Chemical Virology

Decorating the interior of the cowpea chlorotic mottle virus

Een wetenschappelijke proeve op het gebied van de
Natuurwetenschappen, Wiskunde en Informatica

Proefschrift

ter verkrijging van de graad van doctor
aan de Radboud Universiteit Nijmegen
op gezag van de rector magnificus prof. mr. S.C.J.J. Kortmann,
volgens besluit van het college van decanen
in het openbaar te verdedigen op maandag 4 april 2011
om 15.30 uur precies

door

Inge Jeannette Minten
geboren op 11 mei 1983
te Sterksel

Promotores:

Prof. dr. Roeland J.M. Nolte
Prof. dr. Jeroen J.L.M. Cornelissen (Universiteit Twente)

Manuscriptcommissie:

Prof. dr. ir. Jan C.M. van Hest
Prof. dr. ir. David N. Reinhoudt (Universiteit Twente)
Prof. dr. Trevor Douglas (Montana State University)

Paranimfen:

Stijn van Dongen
Loes Ruizendaal

Het onderzoek beschreven in dit proefschrift is uitgevoerd met financiële ondersteuning van de Nederlandse Organisatie voor Wetenschappelijk Onderzoek (NWO)

Druk: Ipskamp Drukkers B.V., Enschede
ISBN: 978-90-9025926-0

"Just as physicists recognize light either as electromagnetic waves or as particulate photons, depending on the context, so biologists can profitably regard viruses both as exceptionally simple microbes and as exceptionally complex chemicals".

- R Dulbecco and HS Ginsberg, 1980. Virology (originally published as a section of Microbiology, 3rd. Edn., by Davis et al., Harper and Row, Hagerstown); p. 855.

Table of contents

Chapter 1

Aim and outline thesis

Nanotechnology is an emerging and promising field of research, which focuses on the design and creation of novel architectures on the nanometer scale. Only recently imaging techniques have come into existence that enable visualization of the nanometer domain. Nanotechnology covers a broad range of subjects, ranging from the development of more efficient solar cells to novel drug delivery systems; it therefore has the potential to have a dramatic impact on society. To create the large diversity of nano-structures currently available, a plethora of different techniques are used, which can be divided into two main approaches: a bottom-up and top-down approach. The latter one creates small structures by starting from larger ones. Photo-lithography is an example of this approach, where parts of a thin film which is patterned by light are selectively removed and partly replaced to generate for example a silicon chip. The bottom-up approach on the other hand does exactly the opposite by creating nano-structures starting from small components. The structuring of these components is directed by molecular self-assembly, i.e. the components are designed such, that molecular recognition phenomena between the components causes them to organize into higher ordered architectures. While scientists have just started to comprehend and use self-assembly to create structures, nature uses self-assembly to create architectures of a mind-boggling complexity. The multitude of examples presented by nature can be used to obtain information about the self-assembly process itself, while at the same time natures building blocks such as proteins and nucleic acids can be employed to create novel synthetic structures. The design and synthesis of these novel structures in turn might provide us with more insight in the way nature functions. Nature, for the major part, uses only four basic building blocks to create biomolecules, i.e. sugars, lipids, amino acids, and nucleotides. From these compounds, larger building blocks, such as proteins or DNA are made, which become the construction

materials for even more complex systems. Relatively simple materials like DNA and short peptides are already commonly employed in the creation of synthetic self assembled nano-structures.[1,2] More recently, also complex structures like viruses have found application as nano-building blocks.[3] Viruses and bacteriophages are arguably the most abundant biological entities on earth,[4] and nature has presented us with a large variety of shapes and sizes to choose from (figure 1).

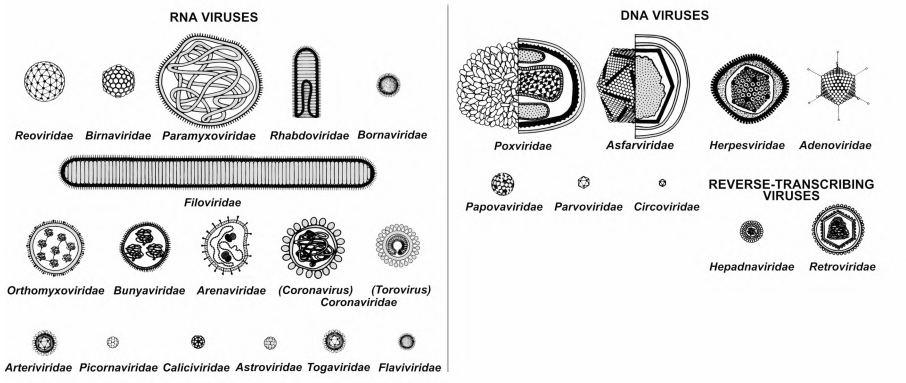

Figure 1 *Cartoon impression of the wide variety of sizes and shapes of viruses.*[5]

Most of these viruses are easily produced in either their natural host or in another host, i.e. a heterologeous expression system, such as yeast or bacteria. Although they are certainly more complex than a single stranded DNA chain or a peptide, viruses have evolved to form relatively simple structures from rather uniform building blocks. Many viruses consist only of a protein shell surrounding a number of nucleic acids. The protein shell is typically composed of multiple copies of one or a small number of different proteins. Viruses usually adopt only one morphology, although some viruses can exist in different shapes depending on the circumstances. The monodispersity of viruses is unparalleled by any synthetic nano-sized structure. This allows for example the growth of very monodisperse nanoparticles on the inside of the virus.[6] Furthermore, the highly symmetric organization of the virus allows precise positional modification of the virus shell which is called capsid. Since the capsid is composed of proteins, it can be functionalized through genetic means or by chemical modification. This property has been elegantly used to create e.g. a light powered water splitting nano-catalyst by Nam et al.[7] They genetically modified the tobacco mosaic virus (TMV) such that it expressed peptides on the capsid exterior that bind to IrO_2, a water oxidizing

catalyst. Through additional chemical modification of the capsid proteins, a photosensitizer was then grafted onto the virus surface, to yield the water splitting system.

Not only the outer surface of a virus is accessible for modification, also by modifying the interface of the constituting proteins, the assembly properties and morphology of the virus can be changed. Modification of the interior presents some additional challenges, such as low accessibility and the presence of the viral genetic material. Often, however, viruses can be produced without their genetic material, resulting in empty capsids. Most viruses contain pores in their capsid shell, through which small molecules can reach the capsid interior. Using this property, a chemotherapeutic agent, taxol, was attached to the interior of the MS2 capsid.[8] While the majority of viral capsids can only be obtained as an intact protein cage, with or without the genetic material, the cowpea chlorotic mottle virus (CCMV) has the special property that after removal of the viral RNA, the capsid proteins dimers can still reversibly assemble, and disassemble, depending on the pH. This property greatly increases versatility of this virus particle.

It is this property of the CCMV that has inspired and enabled the research presented in this thesis. The potential of the well defined protein cage of the CCMV virus to aid the formation of new materials or act as a highly confined reaction space is investigated. In particular the latter is of interest to increase our understanding of (bio)chemical reactions in a spatially restricted environment such as a cell or cellular compartments. By decorating the interior of the CCMV capsid with polymers, micelles, metals or enzymes a function can be added to the capsid, leading to the development of new, responsive materials, delivery systems, and nano-reactors.

In order to place the use of CCMV as a nanoreactor into perspective a literature survey of reactions in protein containers is given in **chapter 2**, together with a short description of the properties of CCMV. In **chapter 3** the encapsulation of a negatively charged redox active polyelectrolyte inside the CCMV capsid is described. Following up on this work, in **chapter 4** it is shown that besides negatively charged polymers also DNA based amphiphile assemblies can be encapsulated. Not only the native capsid protein, but also a genetically modified capsid protein is able to encapsulate negatively charged polymers. This work is reported in **chapter 5**. The expression and purification of the genetically modified capsid protein produced in *Escherichia coli* (*E.coli)* bacteria is also extensively discussed in this chapter. This non-native capsid protein is further modified for the

efficient and controlled encapsulation of enhanced green fluorescent protein (EGFP) (**chapter 6**), while a detailed analysis of the complex assembly mechanism of this modified virus system is described in **chapter 7**. Without the presence of RNA, the capsid is only stable at acidic pH, but in **chapter 8** a method is described to stabilize the capsid at physiological pH, by using the non-native capsid proteins that are appended with a metal binding peptide sequence. Finally, in **chapter 9** the encapsulation of multiple enzymes inside the CCMV capsid is described forming a working protein nanoreactor. In this chapter also the kinetic data obtained from biocatalytic conversions in the protein capsid are discussed in relation to effects of confinement.

References

1. Bromley, E.H.C., et al., *Peptide and Protein Building Blocks for Synthetic Biology: From Programming Biomolecules to Self-Organized Biomolecular Systems.* ACS Chemical Biology, **2008**, *3*, 38.
2. Bath, J. and A.J. Turberfield, *DNA nanomachines.* Nat Nano, **2007**, *2*, 275.
3. Young, M., et al., *Plant viruses as biotemplates for materials and their use in nanotechnology.* Annu. Rev. Phytopathol., **2008**, *46*, 361.
4. Suttle, C.A., *Viruses in the sea.* Nature, **2005**, *437*, 356.
5. Availablefrom: http://www.vetmed.ucdavis.edu/viruses/VirusDiagram.html
6. Douglas, T. and M. Young, *Host-guest encapsulation of materials by assembled virus protein cages.* Nature, **1998**, *393*, 152.
7. Nam, Y.S., et al., *Biologically templated photocatalytic nanostructures for sustained light-driven water oxidation.* Nat Nano, **2010**, *5*, 340.
8. Wu, W., et al., *Genome-Free Viral Capsids as Multivalent Carriers for Taxol Delivery.* Angew. Chem. Int. Edit., **2009**, *48*, 9493.

Chapter 2

Literature overview: reactions inside protein cages

Introduction

Reactions in confined space are currently gaining much interest, because they are expected to lead to unique characteristics of the products and to have a drastic impact on reaction kinetics.[1,2] By performing enzymatic reactions in a confined space, scientists are trying to mimic the way in which enzymes react in nature. Compartmentalization occurs inside cells to create small separate locations with a specialized function, in which high concentrations of specific enzymes can be realized, and to achieve control over the order in which the enzymes react. Often the product of one reaction functions as a catalyst or substrate for the following one.[3] This might significantly decrease formation of side-products and increase reaction rates. The study of enzymes in confined spaces is therefore of crucial importance to gain insight in the complex processes of the cell, and for the development of improved catalytic systems, for the synthesis of new products, etc. Many studies concerning the encapsulation of enzymes by using various spherical particles, such as liposomes, polymersomes, and other spherical aggregates, have been reported.[4-7] However, due to their polydisperse nature, reaction conditions vary considerably from particle to particle. Protein cages are very monodisperse, and would therefore be of great use for the study of enzymatic reactions in confined space. The use of protein cages for enzymatic reactions is relatively new and only recently a few examples have been reported.[8-11] The development of methods to achieve encapsulation of enzymes into the confined space of the CCMV capsid and the subsequent study of enzymatic reactions within it, is one of the main goals of the research presented in this thesis.

Protein cages do not only have potential applications in the study of enzymatic reactions. In the past 20 years, the field of biomineralization has

demonstrated that protein cages can also be used for the formation of highly monodisperse nanoparticles. In these processes the protein cage has different functions; it provides a constrained environment providing the proper conditions for the formation of highly monodisperse nanoparticles, it prevents aggregation of the formed nanoparticles, and in many cases it induces the mineralization reaction. It does so by providing the molecular interactions between the organic and inorganic phases that are crucial for the biomineralization process.[12] An example is the well-defined interior cavity of protein cages like ferritin, which is used as a nanoreactor for the synthesis of inorganic nanoparticles.[13] Also viral cages and even bacterial multi-enzyme complexes have been used for this purpose.[12] These systems have a high charge density at certain regions in their inner cavity, which can act as nucleation sites for the mineralization. Using these strategies, monodisperse single crystal nanoparticles can be grown inside capsids and cages. In addition the protein shell itself can be modified to add extra functionality to the nanoparticles.

Protein Cages

In nature, several proteins exist that carry or store metal ions and minerals. These proteins can often be used as hosts for biomineralization reactions. A well known and intensively studied example is the iron-storage protein ferritin, which is discussed below. Other proteins that are also being used for these purposes are Bacterial DNA binding proteins from starved cells (Dps) ferritin and the small heat shock protein (sHsp).

Ferritin

Ferritins are a class of non-haem iron storage proteins that are produced by animals, plants, fungi, and bacteria.[14] In nature, iron is stored within the multi-subunit protein shell as a hydrous ferric oxide nanoparticle.[15] Ferritins can withstand high temperatures (85 °C) and high pH values (8.5-9.0) which makes them attractive for use in supramolecular and organic synthesis.[16]

Iron-free ferritin (apoferritin or holo-ferritin) is a protein complex of approximately 450 kDa,[17] It consists of 24 polypeptide subunits which assemble into a hollow sphere with an outer and inner diameter of 12 nm and 8 nm, respectively (figure 1).[18] Small channels located at the subunit junctions are required for the release of iron and transport of other metal ions and small organic molecules.[19] The protein shell of ferritin has several functions: it acquires Fe^{II},

catalyses its oxidation and induces mineralization within the cavity, which can store up to 4500 iron atoms. This high storage capacity is achieved by sequestering the iron as a compact mineral, which resembles the structure of the mineral ferrihydrite FeO(OH). To perform these functions, the protein shell contains two structurally similar, but mechanistically different protein subunits, the heavy (H) and light (L) chains. The H-type subunit catalyses the oxidation of Fe^{II} to Fe^{III}, while the L-type subunit promotes mineralization inside the protein cage.

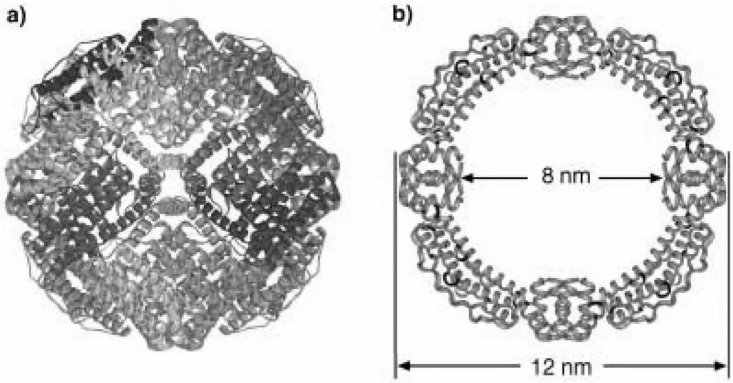

Figure 1 *Schematic representation of horse spleen ferritin a) Complete structure of the 24-meric protein assembly b) Outer and inner diameter of the protein cage* [20]

In 1991, Mann and co-workers pioneered the use of ferritin as a nanosized bioreactor. Their main goal was to produce monodisperse metal particles from other metals than iron inside the ferritin cavity.[21] They applied horse spleen apoferritin (HSFn) to produce iron sulfide particles, as well as manganese oxide and uranyl oxohydroxide crystals. [21-23] The iron sulfide particles were formed by an *in situ* reaction of the native iron oxide cores. By reacting the ferritins with H_2S or Na_2S in an aqueous buffer, the ferrihydrite cores were transformed into iron sulfide nanoparticles. Soon thereafter, they also showed that manganese oxide could be formed by redox-driven reactions in the apoferritin cavity.[22]

Using similar biomineralization strategies, many different inorganic nanoparticles have since then been synthesized inside the core of apoferritin. These include cobalt oxide, cobalt oxohydroxide, chromium hydroxide, nickel hydroxide, indium oxide, cadmium sulfide, cadmium selenide, zinc selenice, magnetite, cobalt/platinum alloys, platinum and other particles.[20,24-29] Some of the recent publications on inorganic nanomaterial synthesis inside the apoferritin cavity are describing in more detail below.

Douglas and co-workers studied the use of the protein constrained iron oxide core of ferritin as a photoreduction catalyst.[30] In this work, the authors showed that the native ferric oxyhydroxide ferrihydrite Fe(O)OH encapsulated within the protein cage of ferritin could act as a semiconductor photocatalyst for the reduction of the highly toxic Cr^{VI} to the more benign Cr^{III}. With this strategy at hand, the same group used the ferritin system to catalyze the photoreduction of Cu^{II} to a colloidal dispersion of Cu^0 with a narrow size distribution.[31] It was also observed that a higher Cu^{II}/ferritin ratio led to larger Cu^0 particle sizes.

A few years later, the group of Watanabe used a similar approach for the development of a size-selective hydrogenation biocatalyst.[20] They encapsulated a Pd nanocluster inside the apoferritin cavity by *in situ* chemical reduction of encased Pd^{II} ions using $NaBH_4$ (figure 2). This hybrid metal catalyst was subsequently used for the hydrogenation of a series of olefins. Ferritin contains eight negatively charged pores which are located at the junctions of three subunits. Substrate charge might therefore influence the diffusion of substrate into the cavity. The effect of substrate charge on catalytic hydrogenation efficiency of the Pd-apoferritin hybrid was therefore evaluated. When anionic or bulky substrates were used, only low turnover frequencies were obtained, but when small, cationic substrates were utilized, high turnover frequencies were observed. These results suggest that the system discriminates based on the size and charge of the substrates.

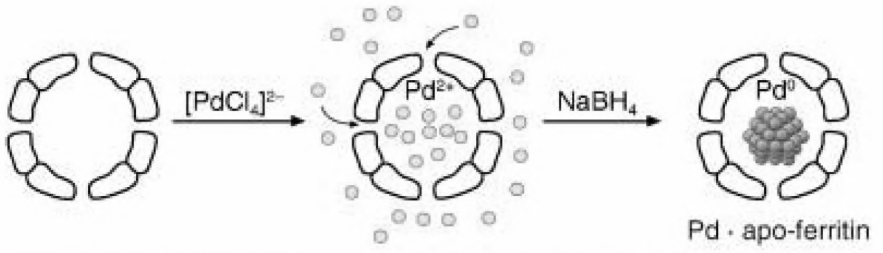

Figure 2 *Schematic representation of the preparation of Pd-apoferritin.*[20]

After having successfully incorporated Pd-complexes in the apoferritin cavity, Watanabe and co-workers proceeded with the encapsulation of Rh^{II}-complexes for the polymerization of phenylacetylene inside the cavity (figure 3).[32] Because of the constrained environment of the ferritin cage, polymers with a narrower weight distribution were obtained than when phenylacetylene was polymerized by the same

Rh-catalyst in bulk. The resulting polymers also have a restricted molecular weight (~13 kg mol^{-1}), and just a few polymer chains can be prepared within one apoferritin cage. This suggests that the polymerization is limited by the available space in the ferritin cavity.

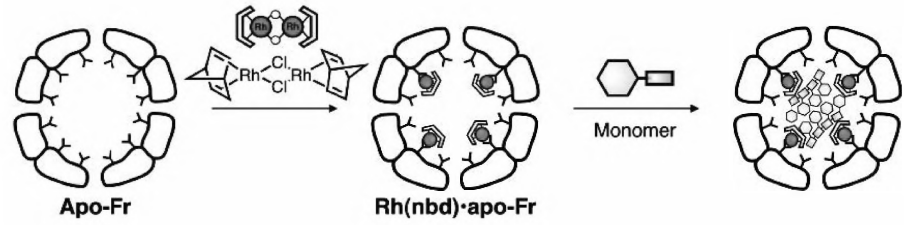

Figure 3 *Schematic representation of a of Rh-catalyzed polymerization inside apoferritin.*[32]

Kramer *et al.*, synthesized silver nanoparticles within the cavity of ferritin by using a similar encapsulation approach combined with genetical modification of the ferritin interior.[33] The authors modified the C-terminus of the L-chain of ferritin with a dodecapeptide, AG4, which was previously identified from a phage display library to be able to reduce silver ions to metallic silver. This method yielded spherical particles with an average diameter of 7 ± 1 nm.

Recently, Kasyutich *et al.*, developed another strategy to synthesize homogeneous silver nanoparticles using ferritin protein cages.[34] It was already suggested in the literature that two parameters are important for obtaining an efficient metal-ion incorporation and a narrow size distribution of the formed nanoparticles inside the protein cage. These are the ratio of external/internal charge distribution[35,36] and the number of available metal binding/nucleation sites.[37] On basis of the literature data the authors utilized ferritin from the hyperthermophilic archaeon *Pyrococcus furiosus* (PfFt), a protein with a distinctive charge distribution and fewer iron nucleation sites compared to other ferritins. By using the PfFt-cage, silver nanoparticles of 2.1 ± 0.4 nm were obtained (figure 4), with a high stability in water solution and a high thermal stability (up to ~90°C). Other ferritins did not give the same results.

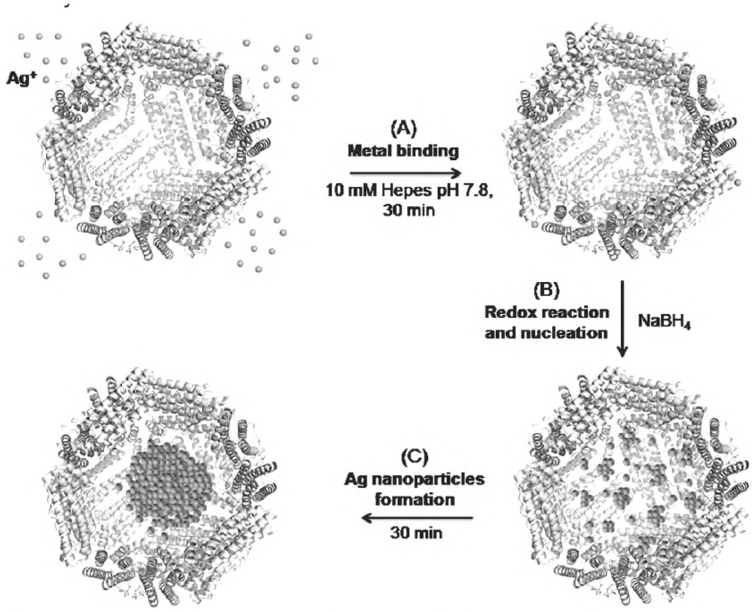

Figure 4 *Silver ions (spheres) access the cavity of Purococcus furiosus and bind to the binding/nucleation centers at the internal surface.*[34]

Aime *et al.* entrapped up to 10 Gd^{III}-chelates in the cavity of apoferritin, which resulted in a complex that exhibited high relaxivity of water protons with potential applications in magnetic resonance imaging (MRI).[38] Though this work is very promising, it does not describe chemical reactions within the cavity of protein cages and is therefore not further discussed here.

The group of Mann also created magnetic nanoparticles by synthesizing ferromagnetic nanocrystals of magnetite (Fe_3O_4) [39,40] and magnetite/maghemite (Fe_3O_4/γ-Fe_2O_3)[41] within the ferritin cavity to yield a magnetic protein, called 'magnetoferritin'. Apoferritin, loaded with various amounts of $iron^{II}$ ions, was oxidized by trimetylamino-*N*-oxide (figure 5). In this method, the authors used the unusual stability of the apoferritin cage at high temperature (60°C) and pH (8.5). This approach resulted in the formation of crystalline inorganic particles with diameters of approximately 6-7 nm. A few years later, Klem *et al.*, reported the preparation of metal oxide nanoparticles containing both Fe and Co oxides.[42] By controlling the addition of Co to the reaction mixture, a two- to fourfold increase in the ferromagnetic blocking temperature (with respect to 'Fe-pure' magnetoferritin), was observed.

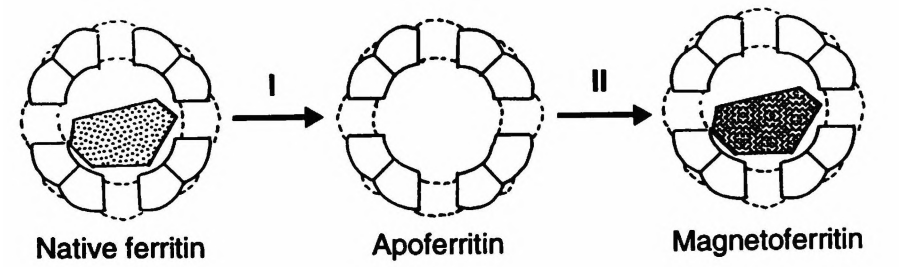

Figure 5 *Schematic representation of the formation of magnetoferritin from native ferritin. First, the native ferrihydrite cores from horse spleen ferritin are removed by dialysis at pH 4.5. Then, apoferritin is reconstituted with a Fe^{II}-solution under slow oxidative conditions at 60°C and pH 8.5.*[40]

Dps ferritin

Dps are a member of the ferritin superfamily. They prevent DNA damage by condensing DNA, and by accumulating iron atoms within their central cavity that might otherwise have produced free hydroxyl radicals by reaction with hydrogen peroxide.[43,44] Dps uses the hydrogen peroxide to oxidize iron, resulting in an iron oxide core similar to that of ferritins.[45] Dps is structurally similar to ferritin. Its 12 Dps subunits have a similar structure as the ferritin subunits and the overall architecture resembles that of the ferritin cage.[46,47] As ferritin contains twice as many subunits than Dps, the Dps cage is smaller, i.e. it has an outer diameter of 8.5 nm and an inner diameter of 5 nm. Structural analysis indicates that the 0.8 nm pores at the subunit interfaces should allow molecular access to the interior cavity of the protein cage.[48] Dps can accumulate 500 iron atoms within its small cavity. It is suggested from X-ray crystal structural data that six clusters of on the interior surface could act as analogues of the mineral nucleation sites in ferritin.[48]

The Dps protein cage was also utilized as a size and shape constrained nanoreactor, as previously described for ferritin. Ferrimagnetic iron oxide,[13] cobalt oxide,[49] cadmium sulfide[50] and platinum nanoparticles[51] were synthesized within this cage. It is worth mentioning that for the first time, the formation of these particles was monitored by Kang *et al.* [51,52] using mass spectrometry

To enhance the formation of nanoparticles, Swift *et al.* replaced the 120 surface accessible hydrophilic residues of Dps with hydrophobic amino acids.[53] Despite these mutations, the Dps proteins were still able to self-assemble and mineralize iron.

Heat shock proteins

Heat shock proteins are produced in high levels in response to cellular stress, and assist in the correct folding of proteins. The sHsp cage from the hyperthermophilic archaeon *Methanococcus jannaschii* consists of 24 subunits of 16.5 kDa which self-assemble into a cage with octahedral symmetry.[54] The major difference between this cage structure and that of the ferritins are the large pores (3 nm in diameter) in the sHsp cage.[55] The cage is stable in the pH range 5-11 and can withstand temperatures up to 70°C. It has en exterior diameter of 12 nm.

Douglas and co-workers functionalized both the interior and exterior surfaces of sHsp with organic and inorganic groups.[56] By genetic manipulation of the protein, thiol groups and endogenous amine groups on both the exterior and interior surfaces were used to attach different molecules for the templating size-constrained synthesis of inorganic materials. Because of the large pores present in the sHsp cage, the functional groups on the interior can easily react with guest molecules. This property has also been used to achieve the chemical attachment of doxorubicin, a chemotherapeutic agent, in the cavity of sHsp.[57]

Like ferritin, sHsp was also used as a reaction vessel for biomimetic mineralization reactions.[56,58] Transition metals, iron oxide and alloy nanoparticles were entrapped and synthesized within the sHsp protein cage.

Recently, the group of Douglas reported the synthesis of a cross-linked, branched polymer network inside the protein cage of sHsp.[59] A genetic sHsp construct with cysteine residues located on its interior surface was used to initiate polymer growth by reaction of the cysteines with a bromo alkyne. Click-chemistry was then applied to synthesize polymers within this cage, by the sequential coupling of multifunctional monomers. It was observed that the stability of cages increased dramatically upon encapsulation of the branched polymer. The cages were able to withstand heat treatment that would otherwise completely disrupt the native protein cage.

Another heat shock protein, Hsp60, which has a cage of 17 nm in diameter, self-assembles into an octadecameric double ring cage structure, which stacks into patterned arrays. The subunits were genetically modified to display a histidine sequence on the interior of the cage, creating a region with high affinity for metal ions. The modified cage was used to template the synthesis of Ni-Pd alloy nanoparticles.[60]

Barrel-shaped protein assemblies

Barrel-shaped protein assemblies, such as molecular chaperonins, have also been used to accommodate inorganic nanoparticles. For example, GroEL is a protein assembly of 14 identical subunits, which stack into rings, yielding a barrel-shape structure of 800 KDa with a diameter of 4.5 nm.[61] Its function is to bind denaturated proteins and assist in their folding which takes place in its cavity. Aida and co-workers used GroEL for the encapsulation and triggered release of nanoparticles.[62] However, as far as we know, no examples of synthesis inside the cavity of chaperonins has been reported.

Viruses

In nature, viruses infect host cells in which they replicate and then exit the cell. During these processes, viruses encounter a broad range of chemical environments. To be able to survive the harsh conditions outside cells but release their cargo when inside, viruses have balanced assembly and disassembly characteristics. All viruses package viral nucleic acid, but many can assemble (naturally or manipulated) into viral capsids, devoid of genetic material, allowing scientists to replace the virus' natural cargo with a non-viral cargo compound. This can be accomplished in two different ways: the cargo can be synthesized in a preassembled capsid, or the protein cage can assemble around an existing cargo. The first method, in which the capsid itself acts as a nanoreactor, is used for the biomineralization of inorganic materials inside the capsid. The latter process is being referred to as *encapsidation.*[63]

In contrast to protein cages, viruses exhibit a much larger variety in sizes and shapes (figure 6). When we look at their morphology, a distinction can be made between icosahedral capsids (quasi-spherical structures) and helical capsids (rod-shaped structures). Icosahedral capsids range in size from 18 to 500 nm in diameter, but filamentous or rod shaped viruses like TMV can form rods up to 2 μm in length.[63]

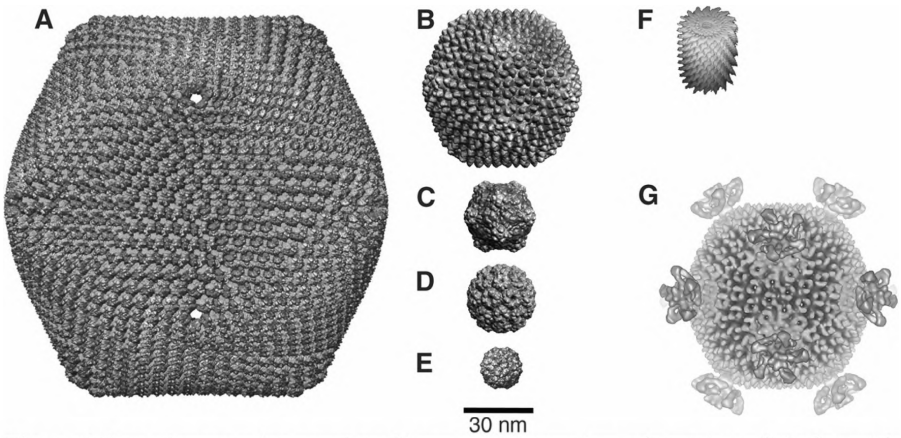

Figure 6 *Cryo-electron micrograph and image reconstructions of a variety of viral capsids.*[64] *A) Paramecium bursaria Chlorella virus type I (PCB-1), 170 nm diameter B) Murine polyoma virus (MPV), 51 nm diameter C) Cowpea mosaic virus (CMV), 31 nm diameter. D) CCMV 28 nm diameter. E) Satellite Tabacco Mosaic Virus (STMV), 18 nm diameter. F) Small section of rod-shaped TMV, 18 by 300 nm. G) Sulfolobus turreted icosahedral virus (STIV) isolated from Yellowstone National Park.*

Virus particles generally consist of several hundred protein molecules which can self assemble to form highly symmetrical structures that contain the viral nucleic acid. The subunits from which the viral capsid is built up can be manipulated both genetically and chemically, often without changing the overall architecture of the virus particles that they form. Like protein cages, the interior of viruses can often be modified via the pores in the capsid. Some viruses can self-assemble in the absence of the endogenous nucleic acids.[12]
The variety in shape and size, and the self-assembly characteristics make viruses very suitable for the controlled synthesis of inorganic materials or the encapsulation of other materials.

When we regard viral capsids as molecular containers, three surfaces can be exploited: the interior, the exterior, and the interface between the protein subunits (Figure 7).[65] In this chapter and in the rest of this thesis, we will focus on the modification of the inside of viral capsids.

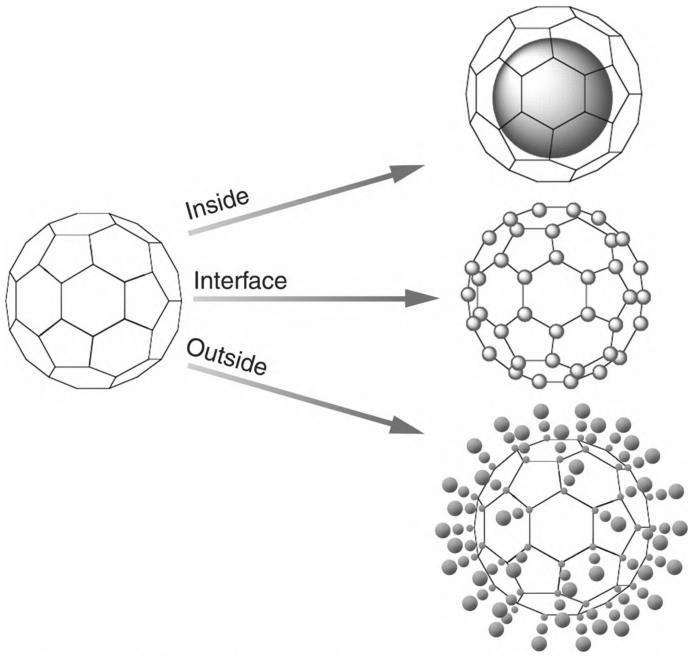

Figure 7 *Schematic representation of the three important interfaces available for chemical and genetic manipulation in a virus particle.*[64] *Herein, we will focus on manipulations of the interior cavity of viral capsids.*

CCMV properties

In this section CCMV and its properties will be briefly discussed. CCMV is the virus used for the studies presented in this thesis. A more detailed description can be found in the thesis of Marta Comellas-Aragonès.[66]

CCMV belongs to the *Bromoviridae* family of plant viruses. It consists of exactly 180 identical capsid proteins, which self-assemble around a central RNA yielding highly defined icosahedral virus particles of 28 nm in size. The virus can be dissociated into protein and RNA, the latter can be removed by precipitation, yielding capsid dimers. All research described in this thesis is conducted with the capsid proteins devoid of viral RNA. The capsid proteins can reassemble into various different structures depending on pH, ionic strength, and the presence of divalent cations in the buffer solution.[67] At pH 5.0 and appropriate ionic strength, the capsid proteins will reassemble into the native 28 nm icosahedral structure, which can be disassembled again into 90 dimers by increasing the pH to 7.5. This process is reversible and can be repeated many times.

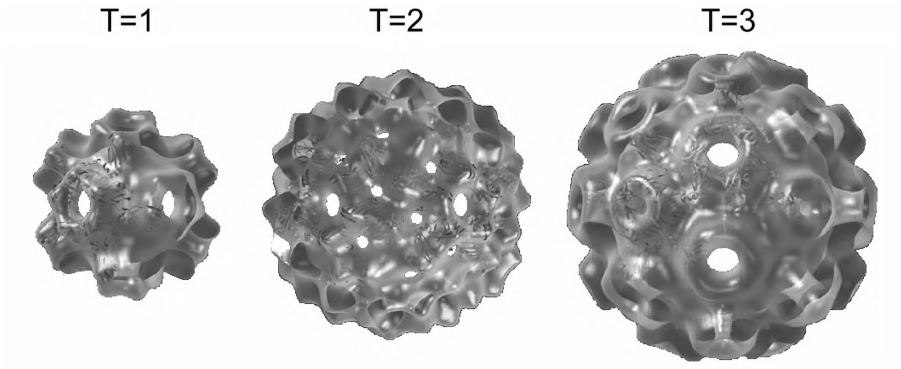

Figure 8 *Structures of T=1, T=2 and T=3 CCMV capsids.*

The 180 identical capsid proteins forming the native capsid are arranged in 12 pentamers and 20 hexamers, forming an icosahedral shell with Caspar-Klug[68] T=3 quasi-symmetry (figure 8).[69] When the N-terminus of the capsid proteins is truncated, the capsid proteins are also known to assemble into T=1 particles, which are 18 nm in size and consist of 60 capsid proteins, or into pseudo T=2 particles of ~22 nm in size, and consisting of 120 capsids.[70] This structure is called pseudo T=2 since geometrically, these particles are not icosahedrons, but dodecahedrons.

The N-termini of the capsid proteins are located at the centers of the hexamers and pentamers, and situated in the capsid interior. The N-termini are highly positively charged and involved in the binding of the negatively charged viral RNA.[71] The C-terminus is highly involved in interdimeric and inter-protein contacts. Capsid proteins with truncated C-termini can not form dimers and fail to assemble into capsids.[72] The C-terminus of one capsid protein of a dimer protrudes in the other capsid protein and is clamped by it.[73] This results in a very stable conformation and therefore the capsid disassembles into dimers, and not into monomers.

Due to its special assembly properties, the CCMV capsid is an exceptionally versatile nano-building block, and its properties will be investigated and employed in the research described in this thesis.

Biomineralization in viruses

Douglas and Young were the first to perform inorganic mineralization reactions inside the CCMV capsid. It was already known from literature[73] that at a pH higher

than 6.5 the capsid swells and 60 pores with a diameter of 2 nm are formed. The swelling is reversed when the pH is lowered to 5.0. This property of CCMV was used for the oligomerization of tungstate (WO_2^{4-}) and vanadate $V_{10}O_{28}^{6-}$ inside the capsid, which occurs when the pH is lowered. This approach yielded particles with an average diameter of 15 nm.[74]

A few years later, Douglas and Young showed that the charge on the protein cage interior of CCMV can be altered from cationic to anionic.[12] This was carried out by replacing nine basic residues at the N-terminus of the CCMV coat protein by glutamic acid residues, but the assembly properties and stability of the formed particle remained similar to the wild-type capsids. The particles that are formed in this manner have a negatively charged interior, which favors interactions with positively charged ferrous and ferric ions. This makes the viral cage suitable for the hydrolysis of Fe^{II} to iron oxide, which leads to the size and shape constrained formation of nearly monodisperse iron oxide particles, with an average diameter of 24 nm.[12]

In a similar approach, the group of Douglas was able to use the CCMV capsid for the biomineralization of TiO_2 nanoparticles.[75] In this study, the authors made use of the wild-type capsid to attract negatively charged Ti^{IV} salts, which were subsequently converted to monodisperse TiO_2 nanoparticles by changing the pH of the solution.

De la Escosura *et al.*, were able to synthesize Prussian Blue nanoparticles with an average diameter of 18 ± 1.7 nm, within the cavity of CCMV.[76] Prussian Blue nanoparticles were formed by a photo-initiated reaction. Negatively charged $[Fe(C_2O_4)_3]^{3-}$ was accumulated in the capsid interior and produced Fe^{II} ions after photoreduction, which further reacted with encapsulated $[FeCN_6]^{3-}$ to form the Prussian Blue clusters (figure 9). This is one of the first examples of the synthesis of magnetic nanoparticles inside viral capsids.

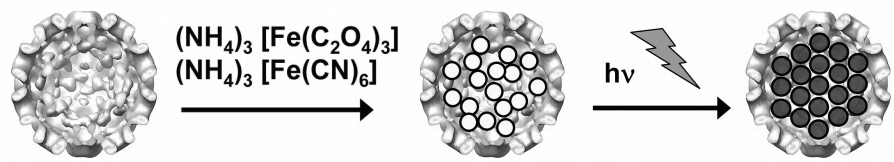

Figure 9 *Schematic representation of the synthesis of Prussian Blue nanoparticles within the cavity of CCMV.*[76]

Another viral capsid that is being used for the constrained synthesis of magnetic nanoparticles is bacteriophage T7.[77] The interior of this virus can be very easily modified because the capsid is formed first, after which the nucleic acids are condensed within it. The empty capsid is therefore also stable without DNA. The T7 phage has an outer diameter of 55 nm, and the interior diameter is around 40 nm. The growth of cobalt particles inside the T7 phage was achieved by incubating the empty capsids with a Co^{II} solution, after which the mixture was reduced by sodium borohydride to yield cobalt nanoparticles. By using this approach, uniform cobalt particles of 42 ± 2 nm were obtained.

Bacteriophage MS2 is an icosahedral capsid consisting of 180 identical subunits, which self-assemble to form particles with a diameter of 27 nm. An important characteristic of this viral capsid is that it has 32 pores of about 1.8 nm, which enable the diffusion of small molecules into and out of the capsid interior. The group of Francis altered the interior of the MS2 capsid by functionalizing the tyrosine residues using diazo-transfer and hetero Diels-Alder reactions.[78] This modification was used to confine Gd^{3+} binding ligands in the capsids, which could be used for MRI contrast imaging.[79] Such chemical alterations of the interior of the capsid may also facilitate chemical or enzymatic reactions in the confined space of the nanoreactor, however, such reactions have not been carried out yet.

TMV is a helical, rod-shaped virus of which the interior has been used for the synthesis of nanowires, due to its anisotropic shape and stability. TMV is stable in a wide range of conditions; it can withstand pH ranges from 3.5 to 9.0, temperatures up to 90^0C and organic solvents.[65] The TMV capsid is made up of 2130 identical subunits, which are recruited by the RNA and assemble around it into a 300 x 18 nm helical structure, with a core of 4 nm in diameter (figure 10). The structure of TMV can be seen as a series of rings (comprising 14 subunits) stacked on top of each other. This results in an overall chiral structure and inherent asymmetry which does not exist for icosahedral viruses. This property allows for chemical and physical differentiation on one end of the helical rod. TMV can easily be assembled in vitro and the TMV aspect ratio is determined by pH, ionic strength and protein concentration.[16] TMV can also be produced in very high quantities (kilogram scale) from infected plant material and because of its high stability, it can serve as a template in a wide range of conditions.

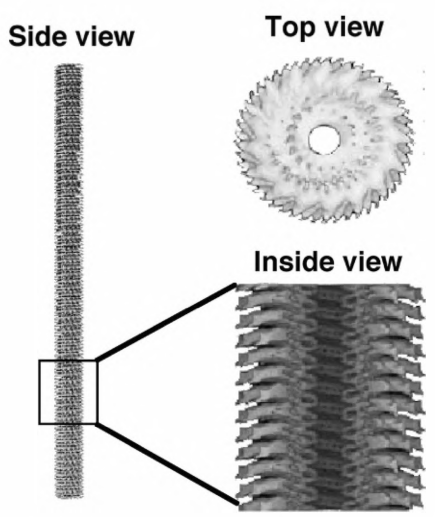

Figure 10 *Cryo-electron imaging reconstructions of tobacco mosaic virus (TMV). The virus is 300 nm in length, has a diameter of 18 nm and a 4 nm channel.*[16]

The channel of TMV was first used for the synthesis of Ni and Co nanowires with lengths up to micrometer range,[80,81] suggesting end-to-end assembly of individual capsids. Later, the narrow channel was used to grow copper nanowires of 150 nm in length and 3 nm diameter[82] and bimetallic alloys of CoPt, $CoPt_3$ and $FePt_3$ up to 100 nm length and 4 nm diameter.[83]

Furthermore, formation of small silver nanoparticles by photochemical reduction of Ag^I salts at pH 7 within the channel of TMV has also been reported.[84] In the same publication the authors describe the use of the exterior surface of TMV for the controlled formation of gold and platinum nanoparticles. The group of Francis reported a method for site-selective modification of both the interior and exterior surfaces of TMV to expand the range of applications for which the capsid can be used.[85]

The exterior surface of TMV has been widely used as template for the synthesis of nanoparticles and nanotubes.[81,84] These nanotubes are grown from deposited clusters of palladium, platinum, and gold on the exterior TMV surface.

The M13 bacteriophage is another helical virus, which has been extensively studied and has proven to be a valuable tool for the development of functional nano-materials. The M13 bacteriophage does not have an inner cavity in which reactions can be performed, but although this literature overview focuses on

reactions on the interior of protein cages, the research in the group of Prof. Belcher is of such great interest that it is briefly discussed here.

The M13 bacteriophage consists of 2700 copies of the p8 coat protein that assemble around circular single stranded DNA, and minor coat proteins which cap the ends of the structure. The coat proteins can be genetically modified to display peptides with metal-binding sequences. By modification of the minor coat proteins situated at the virus ends with ZnS binding sequences, and incubation with a ZnS liquid crystalline suspension, the viruses form highly ordered composite materials.[86] A similar strategy was used to position the virus on single-walled carbon nanotubes (SWNT). Besides modification of the minor coat proteins with SWNT binding peptides, the major coat proteins of this virus were modified with peptides capable of nucleating amorphous iron phosphate (a-$FePO_4$). A crosslinked network of a-$FePO_4$ coated viruses attached to SWNTs was thus created (figure 11). This network is highly conductive, and was used as the positive electrode of a lithium-ion battery.[87] This battery is very light-weight and can be produced and disposed of in environmentally friendly manners, in contrast to conventional Li-ion batteries.

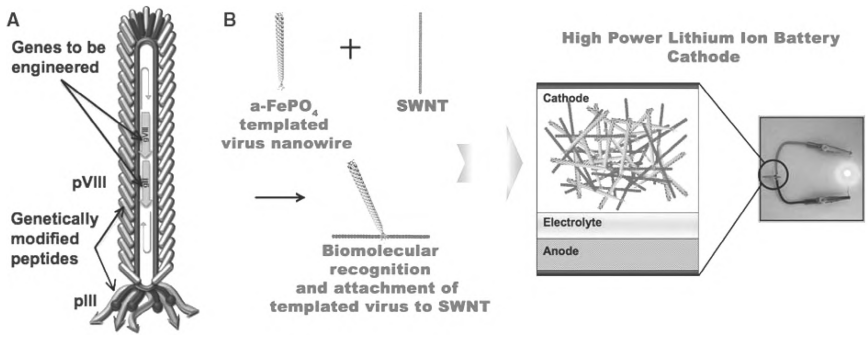

Figure 11 *Schematic representation of the formation of a virus based battery. A) The minor coat proteins of the M13 bacteriophage are modified with SWNT binding peptides, and the major coat proteins are modified with peptides capable of nucleating a-$FePO_4$. B) Addition of the a-$FePO_4$ templated virus nanowires to SWNTs results in a highly conductive network, which is used as the positive electrode of a lithium ion battery.* [87]

Enzymatic reactions inside viral capsids

In the past few years, there has been a growing interest in mimicking cellular compartments. As mentioned in the introduction, viruses could serve as useful model systems for the study of enzymatic reactions in confined spaces, due to their monodispersity and small internal volume. In addition many viruses contain pores through which substrates can enter the capsid. In order to construct a viral nanoreactor, their assembly and disassembly properties could be of use, as well as the accessibility of the capsid to genetic and chemical modification. Despite these promising applications, however, only very few of these viral nanoreactors have been produced.[8,88,89] One of the main goals of the research presented in this thesis is therefore to construct a viral nanoreactor and measure the effect of confinement on enzymatic activity.

The CCMV capsid is very suitable for the encapsulation of enzymes, because under influence of the pH it can disassemble into its dimers, and reassemble again. When enzymes are added to the protein dimers and the pH is lowered to 5.0, enzymes are entrapped inside the capsid (figure 12).

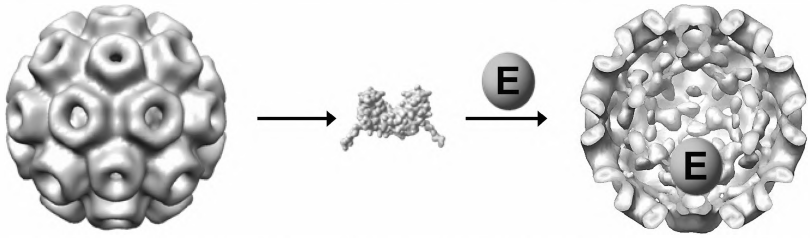

Figure 12 *Schematic representation of enzyme encapsulation in the CCMV capsid. After disassembling into dimers at pH 7.5, the guest enzyme is added and upon decreasing the pH to 5.0 the capsid assembles with the enzyme entrapped in its cavity.*[8]

Comellas-Aragones *et. al* used this approach to incorporate horseradish peroxidase (HRP) enzymes in the inner cavity of CCMV.[8] The goal of this project was to study the activity of one single-enzyme present in the capsid. Therefore, the concentration of the enzyme in buffer was lowered to make sure that only one enzyme or no enzyme was encapsulated in the capsid. The enzymatic activity of the encapsulated HRP was examined by confocal fluorescence microscopy. The non-fluorescent dihydrorhodamine 6G was used as a substrate. Under influence of HRP this substrate is oxidized to rhodamine 6G which is highly fluorescent and thus can easily be monitored. The experiment showed that the capsid acts as a nanoreactor and is permeable for both substrate and product (figure 13).

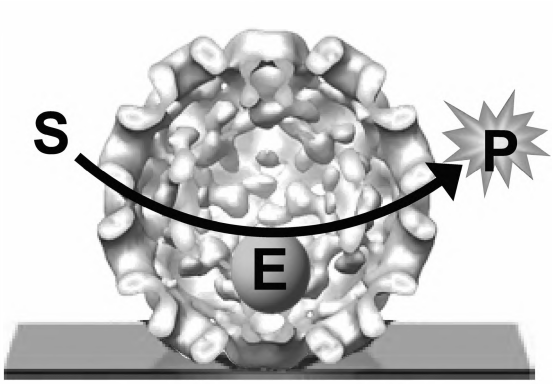

Figure 13 *Single enzyme study on HRP (E) encapsulated within CCMV. The substrate (S) diffuses into the capsid and is converted into a fluorescent product (P) which accumulates before it diffuses out of the capsid.*

Not only plantviruses, but also animal viruses, like the polyomavirus family have been engineered in order to obtain nanocapsules. Simian virus 40 (SV40) and murine polyomavirus (MPV) have been studied extensively for their potential use in delivering biologically active materials.[90] These are icosahedral, non-envelopped DNA viruses, with an outer diameter of about 45 nm. Their capsid consists of 72 pentamers of the major coat protein VP1 (figure 14).[91,92] The minor coat proteins VP2 and VP3 are associated with the VP1 proteins and are located on the capsid interior. Recombinant VP1 pentamers can self-assemble into genome-free icosahedral virus-like particles (VLPs) both *in vitro* and *in vivo.* [93-95]

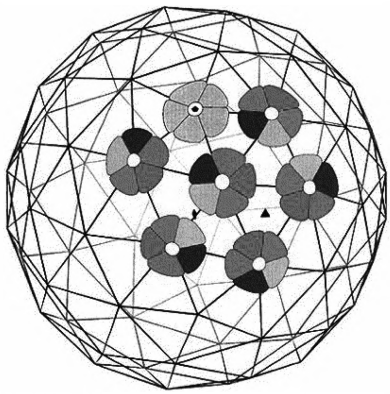

Figure 14 *Architecture of the SV40 capsid. Arrangement of the pentamers on the T=7 icosahedral lattice.*[92]

Recently, many VLPs have been exploited as delivery vehicles.[90] For polyomaviruses this has been achieved by genetically fusing heterologous proteins to VP1 proteins.[96,97] However, the particles formed in these studies were irregular and had showed poor solubility when expressed in *E. Coli.*

Inoue *et al.* reported a novel strategy for the encapsulation of proteins and enzymes into SV40 VLPs.[88] In this work, EGFP was fused to the C-terminus of the VP2 and VP3 minor coat proteins, using a flexible linker. When these constructs were co-expressed with VP1, VLPs with EGFP inside were obtained. The resulting capsids retained their ability to bind to, and enter cells, after which the EGFP was delivered. Next, yeast cytosine deaminase (yCD) was encapsulated via the same method. This enzyme converts cytosine to uracil and is used as a prodrug-converting enzyme in gene therapy. It was demonstrated that upon encapsulation the yCD enzymes retained most of their activity.

To the best of our knowledge, no enzymatic reactions inside other capsids have been described so far.

Helical viruses have not been used for the encapsulation of enzymes because their internal channel is too narrow for these purposes.[1] However, Candida Antarctica Lipase B (CalB), more recently renamed Pseudozyma antarctica lipase B (PalB), has been linked to the surface of Potato Virus X (PVX).[98] This virus particle has a flexible rod-shaped structure, which is 500 nm in length and has a diameter of 13 nm. It is build up from 1270 identical coat proteins. The authors linked the lipase to the N-terminus of the coat protein by constructing a PalB-coat protein fusion protein. In this way the PalB was positioned at the exterior of the virus particle (figure 15). The lipase remained catalytically active, and this work suggests that in the future multiple types of enzymes can be linked to the exterior of a viral capsid, in order to catalyze multi-step enzymatic synthesis. These types of reactions are not constrained by the dimensions of the virus, and will therefore not be discussed further in this thesis.

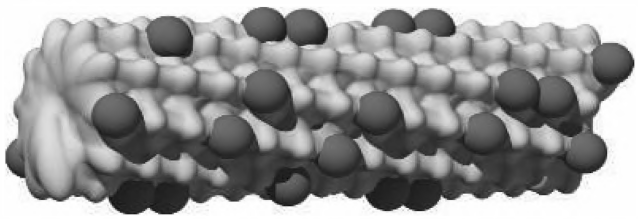

Figure 15 *Schematic representation of the PVX virus particle (grey) with the PalB enzyme molecules (dark grey) on its surface.*[98]

Molecular microcompartments and bacterial enzymes

A family of bacterial proteins that resemble virus capsids are the encapsulins (in literature also referred to as linocin-like proteins), which form assemblies with a diameter of 20 nm and have icosahedral symmetry. Although the sequence similarity between encapsulins and viral proteins is weak, the proteins show great structural similarity to virus capsid proteins. These nanocompartments package enzymes, and are thus excellent examples of nanoreactors. The group of Ban studied this protein family[99] and found two different proteins that can be packaged within the encapsulins; i.e. a peroxidase or a protein that has the same fold as ferritin monomers and contains the ferroxidase active site. Interestingly, the pores of encapsulin resemble those of ferritin. The proteins are attached to the encapsulin shell using a C-terminal extension that interacts with a binding site of the encapsulin proteins.

Besides nanocompartments, many bacterial species produce microcompartments to confine enzymes that have toxic or volatile intermediates, or to increase the local enzyme concentration. These microcompartments are solely composed of proteins. The carboxysome for example is a bacterial microcompartment that packages the enzyme ribulose bis-phosphate carboxylase oxygenase (RuBisCO) by forming polyhedral bodies with a diameter of 80–140 nm.[100-102] Other examples are the Eut and Pdu organelles, which also are likely to package enzymes from metabolic pathways.[103,104] An eukaryotic microcompartment built of protein subunits is the vault particle.[105]

In contrast to protein microcompartments, where structural proteins form the walls of the compartment, multi-enzyme complexes exist, in which the enzymes themselves form the walls of the complex. Substrates can enter these complexes but they are not permeable to other macromolecules. Examples are lumazine synthase,[106] fungal fatty acid synthase[107] and the pyruvate dehydrogenase complex.[108]

Bacterial microcompartments

From diverse bacteria it is known that they produce proteinaceous intracellular compartments, to entrap enzymes to confine them and increase their local concentration.[99] However, the mechanisms by which bacterial microcompartments like Ethanolamine utilization (Eut) and Propanediol utilization (Pdu) encapsulate their enzymes is not very well understood.

The best studied bacterial microcompartment is the carboxysome, which confines CO_2 in the vicinity of RuBisCO to enhance autotrophic CO_2 fixation.[11] Although this microcompartment can be viewed as a potential nanoreactor for enzymatic reactions, since in nature its function is already to encapsulate enzymes, no such applications of the carboxysome have been published yet.

Another bacterial microcompartment that gains interest is the Pdu which is used for the B_{12}-dependent degradation of 1,2-propanediol (1,2-PD). The first two steps of 1,2-PD degradation occur in the lumen of Pdu, the last two in the cytoplasm of the cell.[11] Pdu consists of 14 different polypeptides, including seven different shell proteins and four enzymes.[109] The encapsulated enzymes include CoA-dependent propionaldehyde dehydrogenase (PduP), which convert 1,2-PD to propionyl-CoA via a propionaldehyde intermediate. The function of Pdu is to sequester propionaldehyde to protect cells from DNA damage and cytotoxicity.

Fan *et al.* showed that short N-terminal peptides of Pdu and other proteinaceous microcompartments (MCPs) play a major role in the encapsulation of enzymes.[11] In the same publication, the authors show that fusion of the 18-amino acid N-terminal sequence of PduP to green fluorescent protein (GFP), Glutathione S-transferase (GST) or a maltose-binding protein (MBP) resulted in their encapsulation in Pdu. Though the exact binding mechanism of the N-terminal sequences to the microcompartments is not known, from these experiments it is clear that they are essential for the binding of certain enzymes to the interior of the microcompartments.

Since recombinant enzymes containing these N-terminal extensions are readily made, this discovery offers great potential for the utilization of microcompartments as nanoreactors. As these compartmental structures have only recently been discovered and studied, this has not been accomplished yet, at least to the best of our knowledge. It can be expected that these microcompartments will be used in the future as nanoreactors.

Vault nanoparticle

Vaults are large ribonucleoprotein capsules with dimensions of 42x42x75 nm,[110] with a thin protein shell of 2 nm thick, surrounding an inner cavity large enough to encapsulate hundreds of proteins (figure 16).[9] Vaults consist of four components: 96 copies of the major vault protein (MVP), vault poly(ADP ribose) polymerase (VPARP), telomerase-associated protein 1 (TEP1), and untranslated vault RNA.

Recombinant MVPs expressed in insect cells can self-assemble into a vault-like particle, without the remaining components.

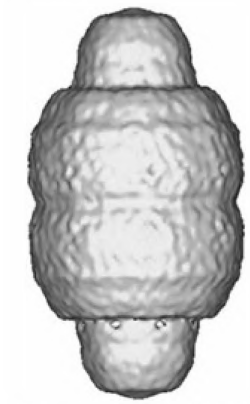

Figure 16 *Vault reconstruction with imposed eight-fold symmetry*[9]

The group of Rome has studied the properties of the vault capsules, mainly for their potential as a drug vehicle.[111,112] Because vaults are abundant in most eukaryotes, including humans, they are less to elicit an immunogenic response than virus capsids or bacterial microcompartments.

First, the authors demonstrated that luciferase could be encapsulated in the cavity of vault, by fusing the luciferase encoding DNA to the vault interaction domain and inserting it into the baculovirus expression vector.[113] In a similar way, a variant of GFP (green lantern) was also sequestered within the vault cavity. Luciferase was still active after encapsulation but its activity seemed to be affected by a diffusion barrier for charged molecules.

A few years later the same group showed that a semiconductor polymer could be encapsulated in recombinant vaults.[114] This study provided important insight in the encapsulation process of nonbiological cargos which could be applied in the sequestration of drug molecules. Recently, gold nanoclusters and histidine-tagged proteins have been targeted to the inner wall of vault particles by using a carrier derived from a vault lumen-associated protein.[111]

Lumazine synthase

Lumazine synthase is a hollow 1MDa bacterial enzyme complex, consisting of 60 subunits. It is involved in the synthesis of lumazine, which is a precursor to riboflavin. X-ray structure analysis showed that lumazine synthase is a hollow

sphere with T=1 icosahedral symmetry, with an outer and inner diameter of 14.7 nm and 7.8 nm respectively.[115] The inner wall of the capsid is negatively charged, due to the presence of glutamic acid residues. The interior of the lumazine synthase capsid can be accessed via hydrophilic channels, which are lined with the glutamic acid residues and located at the ten threefold axes.[106] Channels consisting of symmetrically arranged α-helices are present along the six five-fold pentamer axes[115] (figure 17). The pores formed by these channels at the interior surfaces are surrounded by polar residues and probably play an important role in substrate import and product export from the enzyme catalytic sites located at the interior of the complex.[116] The capsid is highly stable at pH 7 in presence of phosphate buffer or substrate analogues. In absence of ligands, the structure can dynamically be rearranged to form larger particles, with an external diameter of 30 nm.[117]

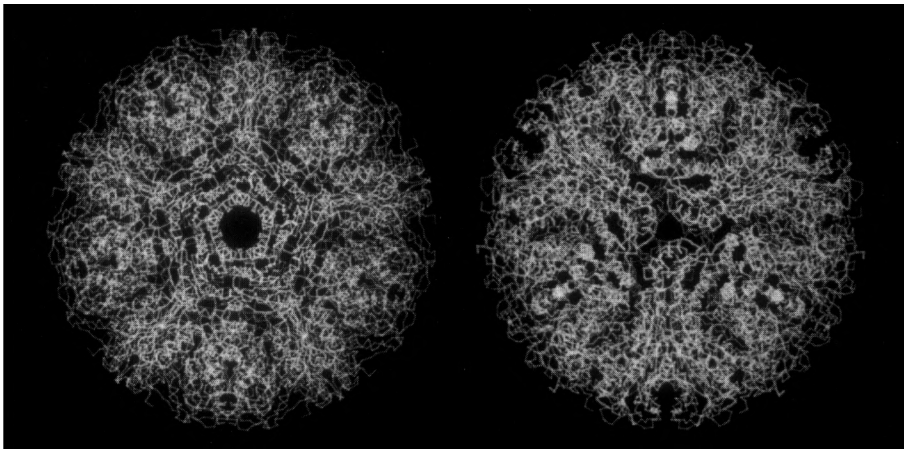

Figure 17 *models of the 60-subunit lumazine synthase capsid. Left: view from the 5-fold icosahedral axis; one of the 12 channels is clearly visible. Right: view from the 3-fold icosahedral axis.*[115]

Shenton *et al*. investigated the potential of the lumazine synthase capsid as bio-nanoreactor for the mineralization of iron oxide.[118] In the environment which the authors used for this reaction (pH 6.5 and absence of ligands) the capsids had an internal diameter of 20 nm and an external diameter of 30 nm, as confirmed by transmission electron microscopy (TEM). Mineralization of iron oxide was induced by adding aliquots of Fe^{II} to the buffer solution. The formation of electron dense iron-containing nanoparticles was confirmed by TEM images and corresponding

energy dispersive X-ray analysis (EDX). The ratio Fe^{II} ions/subunits lumazine synthase was varied in the experiments. More Fe^{II} ions per subunit resulted in more electron dense and more crystalline iron-containing nanoparticles. The diameter of these particles varied between 10 and 15 nm. TEM measurements on the iron-containing capsids revealed that the formed capsids were mostly 30 nm in diameter, but that a small percentage was in the native form of 15 nm in diameter. This suggests that the native form is destabilized by the presence of Fe^{III} ions and transforms into a higher ordered structure. The authors thus showed that it is possible to use a hollow bacterial enzyme complex for the mineralization of inorganic nanoparticles. They also suggested that due to the presence of both positively and negatively charged channels in the capsid, it should be possible to sequester cationic and anionic reactions for the synthesis of other nanoparticles (as previously demonstrated for ferritin). This has not been reported yet.

A few years later, Seebeck *et al.* reported an encapsulation method for lumazine synthase which utilizes a tagging system based on charge complementarity.[10] The authors used lumazine synthase from *Aquifex aeolicus* which they genetically engineered in such a way that four residues per monomer, which project into the lumen, were mutated to glutamic acid residues, to create an additional negative charge on the interior surface of the capsid (figure 18). As the capsids contain either 60 or 180 subunits, these mutations, therefore, could yield 240 or 720 extra negative charges, respectively. Furthermore, a histidine-tag was incorporated at the C-terminus of the protein to facilitate purification. The negatively charged lumazine synthase was brought to expression in *E.coli* and this yielded capsids which can withstand temperatures up to 95°C. Next, the authors showed that a positively charged stretch of amino acids added to a cargo protein led to encapsulation of the cargo protein. This was carried out by fusing a deca-arginine (R_{10}) tag to the C-terminus of GFP. When co-produced with the lumazine synthetase, GFP-R_{10} was encapsulated in the negatively charged lumazine synthase capsid. Sedimentation equibrilium data and scanning force microscopy (SFM) revealed that both T=1 capsids consisting of 60 subunits and T=3 capsids, consisting of 180 subunits were formed.

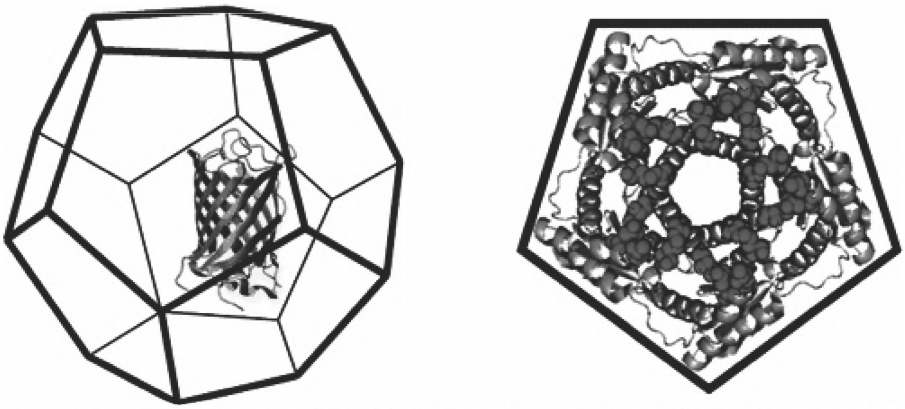

Figure 18 *Left: sequestration of GFP in the lumazine synthase capsid, based on the crystal structure of the T=1 capsid formed by A. aeolicus. Right: lumazine synthase pentamer. In black the four mutated glutamic acid residues.*[10]

Pyruvate dehydrogenase complex

Pyruvate dehydrogenase (PDH) multienzyme complexes of icosahedral symmetry are found in the mitochondria of eukaryotes and gram-positive bacteria.[119] These PDH-complexes contain a dihydrolipoyl acetyltransferase (E2), a pyruvate decarboxylase (E1), a dihydrolipoyl dehydrogenase (E3), and, in the case of the mammalian enzymes, an E3-binding protein and regulatory kinases and phosphatases. The E2 component of this complex consists of a 28 KDa catalytic domain, a 4-kDa peripheral subunit-binding domain (PSBD) and a 9-kDa lipoyl domain, which are connected by long stretches of polypeptide chain thought be largely flexible in conformation.[120] In a native PDH-complex, 60 copies of E2 self-assemble through interactions of their catalytic domains to form an icosahedral structure, to which 42–48 copies of an $\alpha_2\beta_2$-E1 (153 KDa) and 6–12 copies of a homodimeric E3 (100 KDa) bind tightly around the outside (figure 19).[119]

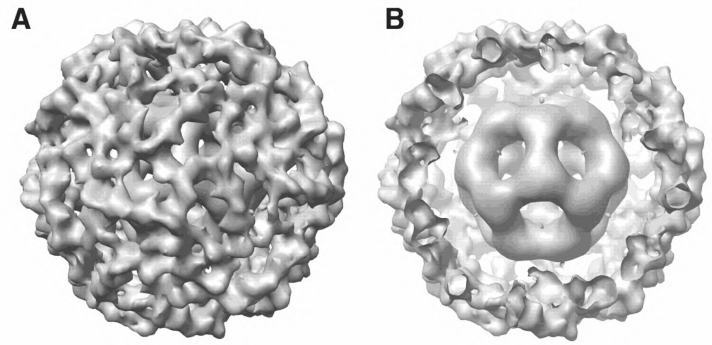

Figure 19 *Schematic representations of the PDH-complex. A) Surface representation of a three-dimensional model of a E2E3 complex, viewed along the 3-fold axis of symmetry. B) Same representation with a portion of the outer protein shell removed to visualize the inner E2-core*[119]

The inner icosahedral shell formed by the E2 subunits is approximately 24 nm in diameter and has 12 openings of 5 nm each. Since this enzyme complex is derived from a thermophilic organism, the association of the subunits into the multi-enzyme complex has been shown to be stable at high temperatures.[121]

Domingo *et al.* showed that it is possible to display antibody epitopes and EGFP on the outer surface of E2.[122,123] This was achieved by replacing the PSDB and lipoyl domains linked to the E2 catalytic domain with either antibody epitopes or EGFP. The resulting fusion protein was able to assemble, thus displaying the epitopes and EGFP on the exterior of the complex. A few years later, Dalmau *et al.* investigated the potential of the E2 scaffold as a potential molecular carrier.[124] In this work, the authors first synthesized a truncated gene encoding for E2 which was optimized for expression in *E.coli.* By mutation of the E2 encoding gene, selected aminoacids were changed to aminoacids with different physicochemical properties. Even after 120 mutations, the E2 complex still assembled into a dodecahedral structure with high thermal stability. The cysteine mutants of the scaffold were used to couple two fluorescent dyes, leading to its sequestration into the inner cavity of the E2-complex. TEM images showed that even after covalent encapsulation of the molecules in the inner cavity, the dodecahedral structure remained intact.

In this work, the authors describe a clear method to encapsulate foreign cargos in the cavity of the E2-complex. Their work can be regarded as a tool for the encapsulation of enzymes in the future. However, no enzymatic or chemical reactions have been reported yet in the constrained environment of the E2-cage.

Concluding remarks

In this chapter many different examples of protein-based nanoreactors have been described. These self-assembled structures all exist in nature and can be chemically or genetically modified for the creation of nanoparticles, or nano catalysts

Protein cages from ferritin and ferritin-like proteins were the first bio-nanoreactors used for biomineralization reactions in a constrained environment. A very wide range of inorganic materials has been successfully synthesized within these cage-structures. A few years later, the use of viral cages for the same purposes became of great interest. Viral cages offer a similar constrained reaction environment but exist in a wider range of shapes and structures, for example the rod-shaped TMV and the icosahedral CCMV, which enables the formation of nanoparticles of more diverse sizes and shapes. Also, some of the viruses are known to be able to change their morphology upon exposure to a changing environment, which increases the possibilities even further. CCMV and TMV are now commonly used to nucleate minerals in their cavities, but this field is still young and it is very probable that in the future more applications for viral capsids will be developed, like MRI contrast imaging agents.

As mentioned before, one of the main advantages of using protein cages for the synthesis of nanoparticles, is the monodispersity of the resulting particles. This is a direct result of the monodispersity of the protein cages themselves, as the nanoparticles grow until they fill the interior of the capsid. It would be expected therefore, that all nanoparticles formed in one type of protein cage would be of the same size. This is however not always the case, as is shown by the different sizes of the nanoparticles formed within the CCMV capsid. Although all nanoparticles of the same type are very monodisperse, there is a considerable size difference between different types of nanoparticles formed within the CCMV capsid, ranging from vanadate particles of 15 nm in diameter,[74] to iron particles of 24 nm in diameter.[12] This can be due to the presence of a 'void space' between the capsid shell and the nanoparticle. It is probable that the size of this space is influenced by the encapsulated material and possibly by the reaction conditions.

The confined space of protein cages has thus far mainly been used for biomineralization reactions. The metal nucleation sites have also been used to immobilize small inorganic catalysist on the interior, which resulted in a nanoreactor that could initiate polymerization of monomers inside the cage.[32] The application of protein cages for the immobilization of enzymes and the study of

enzymatic reactions has great potential, especially since the discovery of bacterial microcompartments show that nature also uses protein cages for the encapsulation of enzymes to create mini-organelles. Thus far only a few groups have published methods to encapsulate proteins in protein cages[10,11,88,113] and even fewer have reported on enzymatic reactions inside the cages,[8] but the future for this field is very promising. A large part of the research described in this thesis is therefore devoted to this subject.

References

1. de la Escosura, A., R.J.M. Nolte, and J. Cornelissen, *Viruses and protein cages as nanocontainers and nanoreactors.* J. Mater. Chem., **2009**, *19*, 2274.
2. Chen, Q., H. Schonherr, and G.J. Vancso, *Block-Copolymer Vesicles as Nanoreactors for Enzymatic Reactions.* Small, **2009**, *5*, 1436.
3. Vriezema, D.M., et al., *Self-assembled nanoreactors.* Chem. Rev., **2005**, *105*, 1445.
4. van Dongen, S.F.M., et al., *A Three-Enzyme Cascade Reaction through Positional Assembly of Enzymes in a Polymersome Nanoreactor.* Chem.-Eur. J., **2009**, *15*, 1107.
5. Walde, P. and S. Ichikawa, *Enzymes inside lipid vesicles: Preparation, reactivity and applications.* Biomolecular Engineering, **2001**, *18*, 143.
6. Kuiper, S.M., et al., *Enzymes containing porous polymersomes as nano reaction vessels for cascade reactions.* Org. Biomol. Chem., **2008**, *6*, 4315.
7. Ha, W., et al., *Self-assembly hollow nanosphere for enzyme encapsulation.* Soft Matter, **2010**, *6*, 1405.
8. Comellas-Aragones, M., et al., *A virus-based single-enzyme nanoreactor.* Nat. Nanotechnol., **2007**, *2*, 635.
9. Kong, L.B., et al., *Structure of the vault, a ubiquitous cellular component.* Structure with Folding & Design, **1999**, *7*, 371.
10. Seebeck, F.P., et al., *A simple tagging system for protein encapsulation.* J. Am. Chem. Soc., **2006**, *128*, 4516.
11. Fan, C.G., et al., *Short N-terminal sequences package proteins into bacterial microcompartments.* P. Natl. Acad. Sci. USA, **2010**, *107*, 7509.
12. Douglas, T., et al., *Protein engineering of a viral cage for constrained nanomaterials synthesis.* Adv. Mater., **2002**, *14*, 415.
13. Allen, M., et al., *Protein cage constrained synthesis of ferrimagnetic iron oxide nanoparticles.* Adv. Mater., **2002**, *14*, 1562.
14. Theil, E.C., *Ferritin: Structure, Gene Regulation, and Cellular Function in Animals, Plants, and Microorganisms.* 1987. p. 289.
15. Chasteen, N.D. and P.M. Harrison, *Mineralization in ferritin: An efficient means of iron storage.* J. Struct. Biol., **1999**, *126*, 182.

16. Uchida, M., et al., *Biological containers: Protein cages as multifunctional nanoplatforms.* Adv. Mater., **2007**, *19*, 1025.
17. Harrison, P.M., et al., *Ferric Oxyhydroxide Core of Ferritin.* Nature, **1967**, *216*, 1188.
18. Harrison, P.M., et al., *Probing Structure-Function Relations in Ferritin and Bacterioferritin.* Adv. Inorg. Chem., **1991**, *36*, 449.
19. Chasteen, N.D., *Ferritin. Uptake, storage, and release of iron*, in *Metal Ions in Biological Systems, Vol 35.* 1998, Marcel Dekker: New York. p. 479.
20. Ueno, T., et al., *Size-selective olefin hydrogenation by a Pd nanocluster provided in an apo-ferritin cage.* Angew. Chem. Int. Edit., **2004**, *43*, 2527.
21. Meldrum, F.C., et al., *Synthesis of Inorganic Nanophase Materials in Supramolecular Protein Cages.* Nature, **1991**, *349*, 684.
22. Mann, S. and F.C. Meldrum, *Controlled Synthesis of Inorganic Materials Using Supramolecular Assemblies.* Adv. Mater., **1991**, *3*, 316.
23. Mann, S., et al., *Crystallization at Inorganic-Organic Interfaces - Biominerals and Biomimetic Synthesis.* Science, **1993**, *261*, 1286.
24. Douglas, T., et al., *Synthesis and Structure of an Iron(Iii) Sulfide-Ferritin Bioinorganic Nanocomposite.* Science, **1995**, *269*, 54.
25. Meldrum, F.C., et al., *Reconstitution of Manganese Oxide Cores in Horse Spleen and Recombinant Ferritins.* J. Inorg. Biochem., **1995**, *58*, 59.
26. Li, M., C. Viravaidya, and S. Mann, *Polymer-mediated synthesis of ferritin-encapsulated inorganic nanoparticles.* Small, **2007**, *3*, 1477.
27. Yoshimura, H., *Protein-assisted nanoparticle synthesis.* Colloids and Surfaces a-Physicochemical and Engineering Aspects, **2006**, *282*, 464.
28. Tsukamoto, R., et al., *Synthesis of Co_3O_4 nanoparticles using the cage-shaped protein, apoferritin.* Bull. Chem. Soc. Jpn., **2005**, *78*, 2075.
29. Okuda, M., et al., *Fabrication of nickel and chromium nanoparticles using the protein cage of apoferritin.* Biotech. Bioeng., **2003**, *84*, 187.
30. Kim, I., et al., *Photochemical reactivity of ferritin for Cr(VI) reduction.* Chem. Mater., **2002**, *14*, 4874.
31. Ensign, D., M. Young, and T. Douglas, *Photocatalytic synthesis of copper colloids from Cu(II) by the ferrihydrite core of ferritin.* Inorg. Chem., **2004**, *43*, 3441.
32. Abe, S., et al., *Polymerization of Phenylacetylene by Rhodium Complexes within a Discrete Space of apo-Ferritin.* J. Am. Chem. Soc., **2009**, *131*, 6958.
33. Kramer, R.M., et al., *Engineered protein cages for nanomaterial synthesis.* J. Am. Chem. Soc., **2004**, *126*, 13282.
34. Kasyutich, O., et al., *Silver Ion Incorporation and Nanoparticle Formation inside the Cavity of Pyrococcus furiosus Ferritin: Structural and Size-Distribution Analyses.* J. Am. Chem. Soc., **2010**, *132*, 3621.
35. Douglas, T. and D.R. Ripoll, *Calculated electrostatic gradients in recombinant human H-chain ferritin.* Protein Sci., **1998**, *7*, 1083.

36. Giorgi, A., et al., *The unusual co-assembly of H- and M-chains in the ferritin molecule from the Antarctic teleosts Trematomus bernacchii and Trematomus newnesi.* Arch. Biochem. Biophys., **2008**, *478*, 69.
37. Ueno, T., et al., *Process of Accumulation of Metal Ions on the Interior Surface of apo-Ferritin: Crystal Structures of a Series of apo-Ferritins Containing Variable Quantities of Pd(II) Ions.* J. Am. Chem. Soc., **2009**, *131*, 5094.
38. Aime, S., L. Frullano, and S.G. Crich, *Compartmentalization of a gadolinium complex in the apoferritin cavity: A route to obtain high relaxivity contrast agents for magnetic resonance imaging.* Angew. Chem. Int. Edit., **2002**, *41*, 1017.
39. Bulte, J.W.M., et al., *Magnetoferritin - Biomineralization as a Novel Molecular Approach in the Design of Iron-Oxide-Based Magnetic-Resonance Contrast Agents.* Investigative Radiology, **1994**, *29*, S214.
40. Meldrum, F.C., B.R. Heywood, and S. Mann, *Magnetoferritin - Invitro Synthesis of a Novel Magnetic Protein.* Science, **1992**, *257*, 522.
41. Wong, K.K.W., et al., *Biomimetic synthesis and characterization of magnetic proteins (magnetoferritin).* Chem. Mater., **1998**, *10*, 279.
42. Klem, M.T., et al., *Synthetic control over magnetic moment and exchange bias in all-oxide materials encapsulated within a spherical protein cage.* J. Am. Chem. Soc., **2007**, *129*, 197.
43. Zhao, G.H., et al., *Iron and hydrogen peroxide detoxification properties of DNA-binding protein from starved cells - A ferritin-like DNA-binding protein of Escherichia coli.* J. Biol. Chem., **2002**, *277*, 27689.
44. Chiancone, E. and P. Ceci, *The multifaceted capacity of Dps proteins to combat bacterial stress conditions: Detoxification of iron and hydrogen peroxide and DNA binding.* Biochim. Biophys. Acta-Gen. Subj., *1800*, 798.
45. Yang, X.K., et al., *Iron oxidation and hydrolysis reactions of a novel ferritin from Listeria innocua.* Biochem. J., **2000**, *349*, 783.
46. Grant, R.A., et al., *The crystal structure of Dps, a ferritin homolog that binds and protects DNA.* Nat Struct Mol Biol, **1998**, *5*, 294.
47. Bozzi, M., et al., *A novel non-heme iron-binding ferritin related to the DNA-binding proteins of the Dps family in Listeria innocua.* J. Biol. Chem., **1997**, *272*, 3259.
48. Ilari, A., et al., *The dodecameric ferritin from Listeria innocua contains a novel intersubunit iron-binding site.* Nat. Struct. Biol., **2000**, *7*, 38.
49. Allen, M., et al., *Constrained synthesis of cobalt oxide nanomaterials in the 12-subunit protein cage from Listeria innocua.* Inorg. Chem., **2003**, *42*, 6300.
50. Iwahori, K., et al., *Cadmium sulfide nanoparticle synthesis in Dps protein from Listeria innocua.* Chem. Mater., **2007**, *19*, 3105.
51. Kang, S., et al., *Monitoring biomimetic platinum nanocluster formation using mass spectrometry and cluster-dependent H-2 production.* Angew. Chem. Int. Edit., **2008**, *47*, 7845.

52. Kang, S., et al., *From Metal Binding to Nanoparticle Formation: Monitoring Biomimetic Iron Oxide Synthesis within Protein Cages using Mass Spectrometry.* Angew. Chem. Int. Edit., **2009**, *48*, 4772.
53. Swift, J., et al., *Design of functional ferritin-like proteins with hydrophobic cavities.* J. Am. Chem. Soc., **2006**, *128*, 6611.
54. Kim, K.K., R. Kim, and S.H. Kim, *Crystal structure of a small heat-shock protein.* Nature, **1998**, *394*, 595.
55. Liepold, L.O., et al., *Structural transitions in Cowpea chlorotic mottle virus (CCMV).* Physical Biology, **2005**, *2*, S166.
56. Flenniken, M.L., et al., *The small heat shock protein cage from Methanococcus jannaschii is a versatile nanoscale platform for genetic and chemical modification.* Nano Letters, **2003**, *3*, 1573.
57. Flenniken, M.L., et al., *Selective attachment and release of a chemotherapeutic agent from the interior of a protein cage architecture.* Chem. Commun., **2005**, 447.
58. Varpness, Z., et al., *Biomimetic synthesis of a H-2 catalyst using a protein cage architecture.* Nano Letters, **2005**, *5*, 2306.
59. Abedin, M.J., et al., *Synthesis of a Cross-Linked Branched Polymer Network in the Interior of a Protein Cage.* J. Am. Chem. Soc., **2009**, *131*, 4346.
60. McMillan, R.A., et al., *A self-assembling protein template for constrained synthesis and patterning of nanoparticle arrays.* J. Am. Chem. Soc., **2005**, *127*, 2800.
61. Braig, K., et al., *The Crystal-Structure of the Bacterial Chaperonin Groel at 2.8-Angstrom.* Nature, **1994**, *371*, 578.
62. Ishii, D., et al., *Chaperonin-mediated stabilization and ATP-triggered release of semiconductor nanoparticles.* Nature, **2003**, *423*, 628.
63. Aniagyei, S.E., et al., *Self-assembly approaches to nanomaterial encapsulation in viral protein cages.* J. Mater. Chem., **2008**, *18*, 3763.
64. Douglas, T. and M. Young, *Viruses: Making friends with old foes.* Science, **2006**, *312*, 873.
65. Young, M., et al., *Plant viruses as biotemplates for materials and their use in nanotechnology.* Annu. Rev. Phytopathol., **2008**, *46*, 361.
66. Thesis of Aragnonès, M.C., Radboud University Nijmegen, 2010.
67. Lavelle, L., et al., *Phase Diagram of Self-assembled Viral Capsid Protein Polymorphs.* J. Phys. Chem. B, **2009**, *113*, 3813.
68. Caspar, D.L.D. and A. Klug, *Physical Principles in Construction of Regular Viruses.* Cold Spring Harbor Symp. Quant. Biol., **1962**, *27*, 1.
69. Johnson, J.E. and J.A. Speir, *Quasi-equivalent viruses: A paradigm for protein assemblies.* J. Mol. Biol., **1997**, *269*, 665.
70. Tang, J.H., et al., *The role of subunit hinges and molecular "switches" in the control of viral capsid polymorphism.* J. Struct. Biol., **2006**, *154*, 59.
71. Vriend, G., B.J.M. Verduin, and M.A. Hemminga, *Role of the N-Terminal Part of the Coat Protein in the Assembly of Cowpea Chlorotic Mottle Virus - a 500 Mhz Proton Nuclear-Magnetic-Resonance Study and Structural Calculations.* J. Mol. Biol., **1986**, *191*, 453.

72. Zhao, X., et al., *In vitro assembly of cowpea chlorotic mottle virus from coat protein expressed in Escherichia coli and in vitro-transcribed viral cDNA.* Virology, **1995**, *207*, 486.
73. Speir, J.A., et al., *Structures of the Native and Swollen Forms of Cowpea Chlorotic Mottle Virus Determined by X-Ray Crystallography and Cryoelectron Microscopy.* Structure, **1995**, *3*, 63.
74. Douglas, T. and M. Young, *Host-guest encapsulation of materials by assembled virus protein cages.* Nature, **1998**, *393*, 152.
75. Klem, M.T., M. Young, and T. Douglas, *Biomimetic synthesis of beta-TiO2 inside a viral capsid.* J. Mater. Chem., **2008**, *18*, 3821.
76. de la Escosura, A., et al., *Viral capsids as templates for the production of monodisperse Prussian blue nanoparticles.* Chem. Commun., **2008**, 1542.
77. Liu, C.M., et al., *Magnetic viruses via nano-capsid templates.* J. Magn. Magn. Mater., **2006**, *302*, 47.
78. Hooker, J.M., E.W. Kovacs, and M.B. Francis, *Interior surface modification of bacteriophage MS2.* J. Am. Chem. Soc., **2004**, *126*, 3718.
79. Hooker, J.M., et al., *Magnetic resonance contrast agents from viral capsid shells: A comparison of exterior and interior cargo strategies.* Nano Letters, **2007**, *7*, 2207.
80. Knez, M., et al., *Biotemplate synthesis of 3-nm nickel and cobalt nanowires.* Nano Letters, **2003**, *3*, 1079.
81. Knez, M., et al., *Spatially selective nucleation of metal clusters on the tobacco mosaic virus.* Adv. Funct. Mater., **2004**, *14*, 116.
82. Balci, S., et al., *Copper nanowires within the central channel of tobacco mosaic virus particles.* Electrochim. Acta, **2006**, *51*, 6251.
83. Tsukamoto, R., et al., *Synthesis of CoPt and FePt3 nanowires using the central channel of tobacco mosaic virus as a biotemplate.* Chem. Mater., **2007**, *19*, 2389.
84. Dujardin, E., et al., *Organization of metallic nanoparticles using tobacco mosaic virus templates.* Nano Letters, **2003**, *3*, 413.
85. Schlick, T.L., et al., *Dual-surface modification of the tobacco mosaic virus.* J. Am. Chem. Soc., **2005**, *127*, 3718.
86. Lee, S.W., et al., *Ordering of quantum dots using genetically engineered viruses.* Science, **2002**, *296*, 892.
87. Lee, Y.J., et al., *Fabricating Genetically Engineered High-Power Lithium-Ion Batteries Using Multiple Virus Genes.* Science, **2009**, *324*, 1051.
88. Inoue, T., et al., *Engineering of SV40-based nano-capsules for delivery of heterologous proteins as fusions with the minor capsid proteins VP2/3.* J. Biotechnol., **2008**, *134*, 181.
89. Minten, I.J., et al., *Controlled encapsulation of multiple proteins in virus capsids.* J Am Chem Soc, **2009**, *131*, 17771.
90. Garcea, R.L. and L. Gissmann, *Virus-like particles as vaccines and vessels for the delivery of small molecules.* Current Opinion in Biotechnology, **2004**, *15*, 513.
91. Liddington, R.C., et al., *Structure of Simian Virus-40 at 3.8-a Resolution.* Nature, **1991**, *354*, 278.

92. Stehle, T., et al., *The structure of simian virus 40 refined at 3.1 angstrom resolution.* Structure, **1996**, *4*, 165.
93. Kanesashi, S.N., et al., *Simian virus 40 VP1 capsid protein forms polymorphic assemblies in vitro.* J. Gen. Virol., **2003**, *84*, 1899.
94. Kawano, M., et al., *The VP2/VP3 minor capsid protein of simian virus 40 promotes the in vitro assembly of the major capsid protein VP1 into particles.* J. Biol. Chem., **2006**, *281*, 10164.
95. Montross, L., et al., *Nuclear Assembly of Polyomavirus Capsids in Insect Cells Expressing the Major Capsid Protein Vp1.* J. Virol., **1991**, *65*, 4991.
96. Abbing, A., et al., *Efficient intracellular delivery of a protein and a low molecular weight substance via recombinant polyomavirus-like particles.* J. Biol. Chem., **2004**, *279*, 27410.
97. Boura, E., et al., *Polyomavirus EGFP-pseudocapsids: Analysis of model particles for introduction of proteins and peptides into mammalian cells.* FEBS Lett., **2005**, *579*, 6549.
98. Carette, N., et al., *A virus-based biocatalyst.* Nat. Nanotechnol., **2007**, *2*, 226.
99. Sutter, M., et al., *Structural basis of enzyme encapsulation into a bacterial nanocompartment.* Nat. Struct. Mol. Biol., **2008**, *15*, 939.
100. Cannon, G.C., et al., *Microcompartments in prokaryotes: Carboxysomes and related polyhedra.* Appl. Environ. Microbiol., **2001**, *67*, 5351.
101. Kerfeld, C.A., et al., *Protein structures forming the shell of primitive bacterial organelles.* Science, **2005**, *309*, 936.
102. Tanaka, S., et al., *Atomic-level models of the bacterial carboxysome shell.* Science, **2008**, *319*, 1083.
103. Havemann, G.D., E.M. Sampson, and T.A. Bobik, *PduA is a shell protein of polyhedral organelles involved in coenzyme B-12-dependent degradation of 1,2-propanediol in Salmonella enterica serovar typhimurium LT2.* J. Bacteriol., **2002**, *184*, 1253.
104. Kofoid, E., et al., *The 17-gene ethanolamine (eut) operon of Salmonella typhimurium encodes five homologues of carboxysome shell proteins.* J. Bacteriol., **1999**, *181*, 5317.
105. Anderson, D.H., et al., *Draft crystal structure of the vault shell at 9-angstrom resolution.* Plos Biology, **2007**, *5*, 2661.
106. Ritsert, K., et al., *Studies on the Lumazine Synthase/Riboflavin Synthase Complex of Bacillus-Subtilis - Crystal-Structure Analysis of Reconstituted, Icosahedral Beta-Subunit Capsids with Bound Substrate-Analog Inhibitor at 2.4 Angstrom Resolution.* J. Mol. Biol., **1995**, *253*, 151.
107. Jenni, S., et al., *Structure of fungal fatty acid synthase and implications for iterative substrate shuttling.* Science, **2007**, *316*, 254.
108. Aevarsson, A., et al., *Crystal structure of 2-oxoisovalerate and dehydrogenase and the architecture of 2-oxo acid dehydrogenase multienzyme complexes.* Nat. Struct. Biol., **1999**, *6*, 785.
109. Bobik, T.A., et al., *The propanediol utilization (pdu) operon of Salmonella enterica serovar typhimurium LT2 includes genes necessary for formation*

of polyhedral organelles involved in coenzyme B-12-dependent 1,2-propanediol degradation. J. Bacteriol., **1999**, *181*, 5967.

110. Kong, L.B., et al., *RNA location and modeling of a WD40 repeat domain within the vault.* Rna-a Publication of the Rna Society, **2000**, *6*, 890.
111. Goldsmith, L.E., et al., *Utilization of a Protein "Shuttle" To Load Vault Nanocapsules with Gold Probes and Proteins.* Acs Nano, **2009**, *3*, 3175.
112. Lai, C.Y., et al., *Vault Nanoparticles Containing an Adenovirus-Derived Membrane Lytic Protein Facilitate Toxin and Gene Transfer.* Acs Nano, **2009**, *3*, 691.
113. Kickhoefer, V.A., et al., *Engineering of vault nanocapsules with enzymatic and fluorescent properties.* P. Natl. Acad. Sci. USA, **2005**, *102*, 4348.
114. Ng, B.C., et al., *Encapsulation of Semiconducting Polymers in Vault Protein Cages.* Nano Letters, **2008**, *8*, 3503.
115. Ladenstein, R., et al., *Heavy Riboflavin Synthase from Bacillus-Subtilis - Crystal-Structure Analysis of the Icosahedral-Beta-60 Capsid at 3.3-a Resolution.* J. Mol. Biol., **1988**, *203*, 1045.
116. Ladenstein, R., A. Bacher, and R. Huber, *Some Observations of a Correlation between the Symmetry of Large Heavy-Atom Complexes and Their Binding-Sites on Proteins.* J. Mol. Biol., **1987**, *195*, 751.
117. Schott, K., et al., *The Lumazine Synthase-Riboflavin Synthase Complex of Bacillus-Subtilis - Crystallization of Reconstituted Icosahedral Beta-Subunit Capsids.* J. Biol. Chem., **1990**, *265*, 12686.
118. Shenton, W., et al., *Synthesis of nanophase iron oxide in lumazine synthase capsids.* Angew. Chem. Int. Edit., **2001**, *40*, 442.
119. Milne, J.L.S., et al., *Molecular structure of a 9-MDa icosahedral pyruvate dehydrogenase subcomplex containing the E2 and E3 enzymes using cryoelectron microscopy.* J. Biol. Chem., **2006**, *281*, 4364.
120. Milne, J.L.S., et al., *Molecular architecture and mechanism of an icosahedral pyruvate dehydrogenase complex: a multifunctional catalytic machine.* Embo J, **2002**, *21*, 5587.
121. Jung, H.I., A. Cooper, and R.N. Perham, *Identification of key amino acid residues in the assembly of enzymes into the pyruvate dehydrogenase complex of Bacillus stearothermophilus: A kinetic and thermodynamic analysis.* Biochemistry, **2002**, *41*, 10446.
122. Domingo, G.J., S. Orru, and R.N. Perham, *Multiple display of peptides and proteins on a macromolecular scaffold derived from a multienzyme complex.* J. Mol. Biol., **2001**, *305*, 259.
123. Domingo, G.J., et al., *Induction of specific T-helper and cytolytic responses to epitopes displayed on a virus-like protein scaffold derived from the pyruvate dehydrogenase multienzyme complex.* Vaccine, **2003**, *21*, 1502.
124. Dalmau, M., et al., *Thermostability and Molecular Encapsulation Within an Engineered Caged Protein Scaffold.* Biotech. Bioeng., **2008**, *101*, 654.

Chapter 3

*CCMV capsid formation induced by a functional negatively charged polymer**

Introduction

Materials composed of nanometer-sized compounds have gained much interest over the past decades with a growing number of practical utilizations.[1] These structures are more and more often constructed from biomolecules like proteins and DNA.[2-9] Their major advantage with respect to non-biological nano-sized compounds is their exceptional ability to self-assemble into precisely well-defined structures. CCMV is an example of such a defined biological system[10] (see chapter 2).

Sikkema et al. have shown that the negatively charged polymer polystyrene sulfonate (PSS) together with capsid dimers at pH 7.5 forms VLPs of 19 nm.[11] These VLPs likely consist of 60 protein subunits instead of 180 and can be described as a T=1 particle. In order to investigate whether the same principle would apply to other negatively charged polymers, and to add functionality to the system, we describe in this chapter the encapsulation in CCMV of a functional negatively charged polyelectrolyte, polyferrocenylsilane (PFS) polyanions (figure 1).[12]

PFS contains redox active ferrocene units in the main chain and can therefore be oxidized and reduced reversibly,[13,14] either by chemical or electrochemical means.[15] In the reduced state, the polyions have a net negative charge due to the negatively charged sulfonate polymer side groups. [1]

* This work was carried out in collaboration with Yujie Ma, Mark Hempenius and Julius Vancso from the Department of Materials Science and Technology of Polymers, University of Twente and has been published in *Org. Biomol. Chem.*, **2009**, 7, 4685

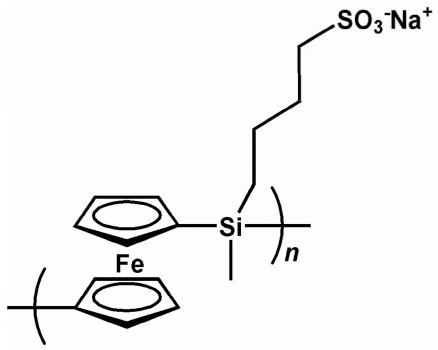

Figure 1 *Structure of the anionic PFS polymer*

Complete oxidation of the polymer will introduce additional positive charges on the ferrocene units in the polymer main chain, which results in a net neutral charge. The formation of the capsid at neutral pH likely depends on the interaction of the protein dimers with a negatively charged species, so switching the polymer towards its neutral state could have an effect on the stability of the capsid. Conversely, the capsid proteins might also have an effect on the redox chemistry of the polymer, since it is known from e.g. the cytochrome family, that proteins can have a significant impact on the redox potential and behavior of a metal center.[16] Here, we report on studies of the interaction of the PFS polyanions with the capsid proteins at neutral pH (figure 2), the redox chemistry of the PFS capsid complex, and the stability of the PFS capsid complex under these circumstances.

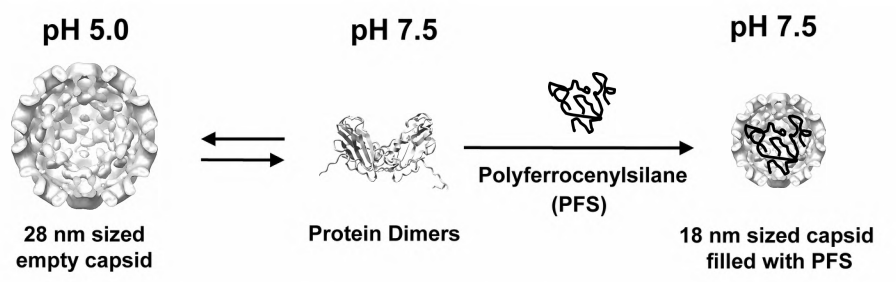

Figure 2 *Schematic representation of PFS encapsulation.*

Results and Discussion

CCMV capsid proteins and PFS polyanions were obtained according to a literature procedure.[14,17,18] The PFS polyanions (figure 1) were dissolved in a small volume of milli-Q (MQ) and added to the capsid protein at pH 7.5 in a ratio of 40 polymer monomeric units per capsid protein, since previous experience with the encapsulation of PSS polymer had indicated that 40 negative charges per capsid monomer was the optimal ratio.[11] The particles were allowed to form for at least 15 minutes at 4°C and were subsequently analyzed by Fast Performance Liquid Chromatography (FPLC) and TEM. The formed particles eluted at V= 1.20 mL as detected at both λ=280 and λ=450 nm (figure 3). The PFS polyanions have a specific absorption at λ=450 nm[19] when it is in the reduced state, whereas proteins do not absorb in this region. This showed that the PFS polyanions co-elute with the protein (figure 3), indicating that it is indeed encapsulated. Non-encapsulated PFS polyanions did not seem to elute from the column, possibly because the PFS polyanions stick to the column.

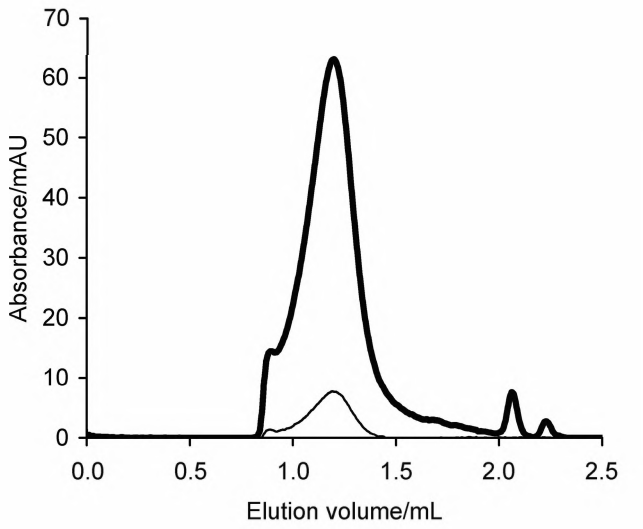

Figure 3 *FPLC trace of capsid proteins mixed with PFS polyanions. The FPLC was equipped with a superose 6 size exclusion column. The bed volume of this column was 2.4 mL. The thick line represents the UV absorption at 280 nm, at which both the protein and the polymer absorb. The thin line represents the UV absorption at 450 nm, which is specific for the PFS polyanions. The elution volume of the peak is 1.20 mL. CCMV capsids of 28 nm in size usually elute at V=1.1 mL, and the capsid dimers usually elute at V=1.8 mL.*

In order to visualize the capsids with TEM, these particles were stained with uranyl acetate, a negative staining agent, revealing capsids with a diameter of d=18 nm in size (figure 4a). This result was confirmed by Dynamic Light Scattering (DLS). This negative staining procedure visualizes the protein material, because the staining agent (uranyl acetate) does not adhere to proteins. In that way the particle is visible as a bright spot, often with a sharp rim, in a darker background. Since the PFS contains iron atoms in the main chain, the PFS-containing particles should also be visible without this staining technique. Indeed, particles of a slightly smaller size than the stained PFS containing capsids were visible without applying staining (figure 4b), exhibiting dark PFS loaded capsules. This is further support of the encapsulation of PFS.

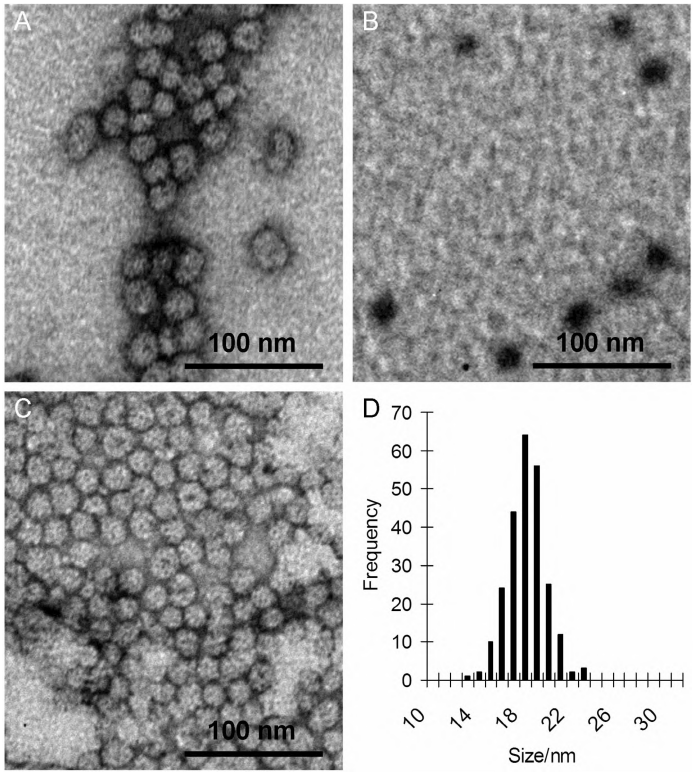

Figure 4 *Size of the PFS-protein nanoparticles as obtained from TEM imaging. A) PFS loaded capsids stained with uranyl acetate. B) Non-stained PFS loaded capsids, (the iron atoms of the PFS polyanions provide the contrast in this case) C) PFS loaded capsids stained with uranyl acetate after dialysis with MQ. D) Size distribution of PFS loaded capsids.*

The percentage of PFS that was encapsulated in the T=1 particles purified by FPLC was determined using the UV absorbance ratio at λ= 280 and 450 nm. The capsid protein absorbs at λ=280 nm, while the PFS absorbs at both these wavelengths and has a 280/450 nm absorbance ratio of 11.67. Since the 280/450 nm ratio of the T=1 particles was consistently 20, the percentage of PFS loading can be calculated, indicating that about 12% of the added PFS is encapsulated. From this data, the concentration of PFS inside the capsids can be determined, yielding a value of 1.9×10^{-2} M PFS, whereas the concentration of the same amount of PFS free in solution would be 5.8×10^{-5} M (see experimental section for calculation details). This means that the concentration of PFS inside the capsids is about 330 times higher than one would expect from statistical encapsulation. This strongly indicates that the encapsulation of PFS is favored, probably because it is needed to stabilize the capsid structure in the absence of the natural RNA.

The PFS containing particles appeared to be quite stable; they survive at least two months of storage at 4°C, and even dialysis in MQ (figure 4c). This is in contrast to the empty T=3 capsids resembling the natural virus particle, which need large amounts of salt to stabilize their structure. This is a strong additional indicator that the PFS polyanions stabilize the T=1 particle, presumably through ionic interactions of the negatively charged PFS polyanions with the positively charged interior of the capsid proteins. Taking these properties into account, it is of interest to study whether the PFS containing particles are still stable when these ionic interactions are altered. Since the PFS polyanion is redox responsive,[14] the polymer can be switched from a negative to a net neutral state, which can potentially disrupt the ionic interactions.

PFS can be oxidized chemically by $FeCl_3$. Unfortunately, this oxidant appeared to harm the capsid proteins. That is, when $FeCl_3$ was added to the capsid protein dimers, the dimer peak in the FPLC disappeared. A new peak at V=2.06 mL appeared, but this is probably due to the oxidized polymer, since a Western blot of the material eluting at this volume revealed that it did not contain any capsid protein. The addition of other oxidizing agents (KI, $KMnO_4$, Cl_2 (g)) gave the same result. $(NH_4)_3$ $[Fe(CN)_6]$ did not appear to have such a detrimental effect on the protein, since the capsid protein dimer was still clearly visible by FPLC. Unfortunately, however, this compound did not oxidize PFS completely, as was visible by the absence of a change in optical appearance. This in contrast to the oxidation with $FeCl_3$, where the optical appearance changed immediately from yellow to dark green/blue, after addition of $FeCl_3$. This finding was confirmed by UV-Vis analysis

of the samples (figure 5). Oxidized PFS has a small absorption band at λ=640 nm. When the PFS loaded capsids were oxidized with $FeCl_3$, the absorption at λ=640 nm was larger than when the PFS loaded capsids were oxidized with $(NH_4)_3$ $[Fe(CN)_6]$. This indicates that $(NH_4)_3$ $[Fe(CN)_6]$ is not as effective in oxidizing the PFS, as $FeCl_3$. Therefore, it seemed that the balance between an oxidizing agent that was sufficiently strong to oxidize the PFS, but not so strong as to oxidize the protein, would be very delicate or even non-existent.

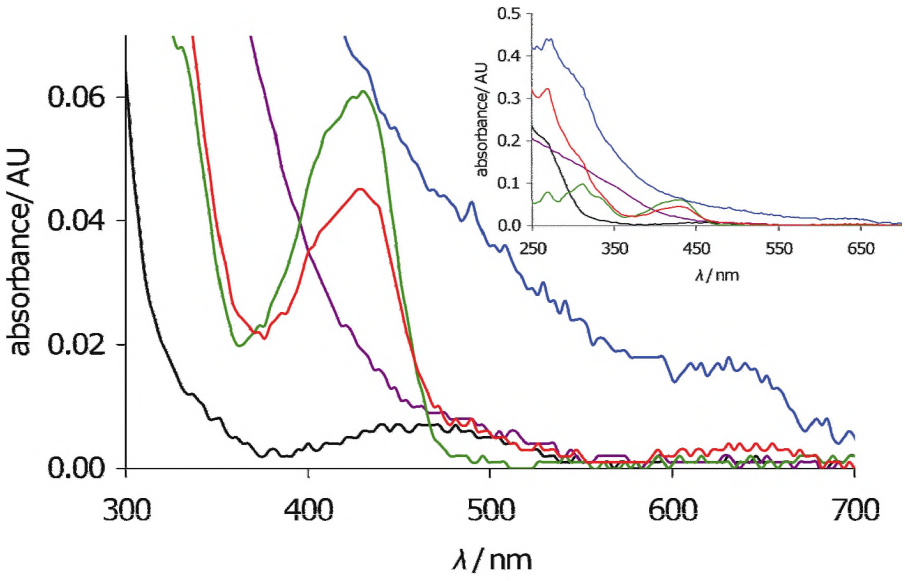

Figure 5 *UV-Vis absorbance spectra of PFS loaded capsids (black line), $FeCl_3$ (purple line), PFS loaded capsids oxidized with $FeCl_3$ (blue line), $(NH_4)_3$ $[Fe(CN)_6]$ (green line) and PFS loaded capsids oxidized with $(NH_4)_3$ $[Fe(CN)_6]$ (red line). Inset: full UV-Vis absorbance spectra.*

For that reason electrochemical oxidation was carried out to oxidize the PFS polyanions. Experimentally, 3 platinum wires acting as the working, reference and counter electrodes respectively were submerged into the T=1 particle containing solution and attached to a potentiostat with a cyclic voltametry (CV) setup. The recorded cyclic voltammograms showed that the electrochemical behavior of the PFS inside the capsid deviated significantly from the behavior of PFS polyanions that are free in solution (figure 6). Normally, PFS can be oxidized and reduced

reversibly, but once encapsulated this did not seem to be the case. In the first cycle, the oxidation peak was about as strong as the oxidation peak of the PFS free in solution, but the reduction peak was much weaker. In the subsequent cycle, the oxidation peak was also weaker, and the reduction peak decreased even further, until in the third cycle the peaks had disappeared completely. This could point to an incomplete reduction of the PFS polyanion. With each successive cycle more PFS polyanions remain oxidized, until after only three cycles, no electrochemical response can be detected anymore. Also the positions of the redox peaks in the voltammograms were shifted upon encapsulation with regard to normal PFS. This could be related to an effect of the protein environment, as proteins are also known to have a profound influence on the redox potentials of for example iron in the cytochrome protein family.[16]

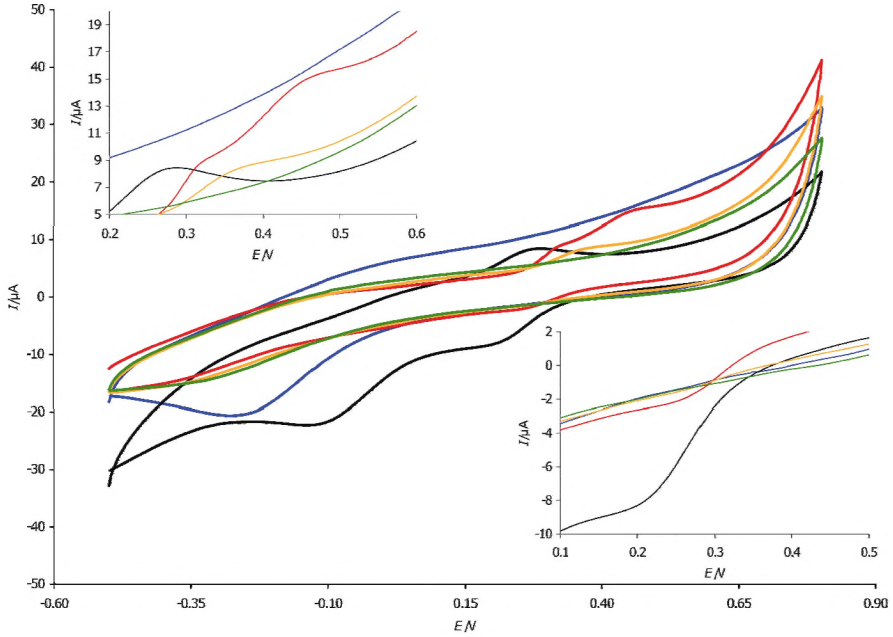

Figure 6 *CV diagrams of PFS loaded capsids. Blue line: buffer, black line: free PFS in solution, red line: the first oxidation cycle of the PFS loaded capsids, orange line: the second oxidation cycle of the PFS loaded capsids, green line: the third oxidation cycle of the PFS loaded capsids. Inset top left corner: enlargement of oxidation peaks in the CV diagram. Inset bottom right corner: enlargement of reduction peaks in the CV diagram.*

In order to check whether the redox response is really caused by the PFS polyanions and not by the protein shell surrounding it, a redox unresponsive polyanion, PSS, was encapsulated and CV was measured in the same way. Indeed, no redox response was observed (figure 7), confirming that the redox response of the PFS loaded capsids was caused by the ferrocene units, and not by the protein.

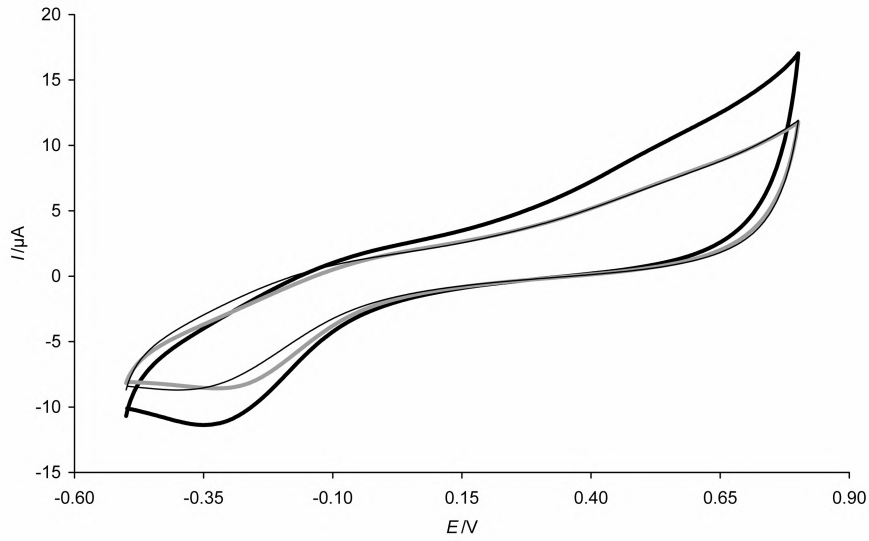

Figure 7 *CV diagrams. Thick black line: buffer, grey line: free PSS in solution, thin black line: PSS loaded capsids.*

To investigate whether the capsid stability was affected by the oxidation and reduction of the polymer, the particles were studied by TEM at three different time points: before oxidation, immediately after oxidation, and following reduction (figure 8). In order to obtain a representative view of the electrochemical process the TEM samples were prepared by dipping the specimen grids in the solution at designed time points. Since the concentration of particles is much higher in these experiments, as compared to the concentration normally used to prepare TEM grids, the images are somewhat blurred. At each time point capsids were still visible, although directly after oxidation the capsids seemed less well-defined. In this case, however, the contours of the PFS containing particles were still clearly visible; therefore it is unlikely that the capsids are completely disassembled.

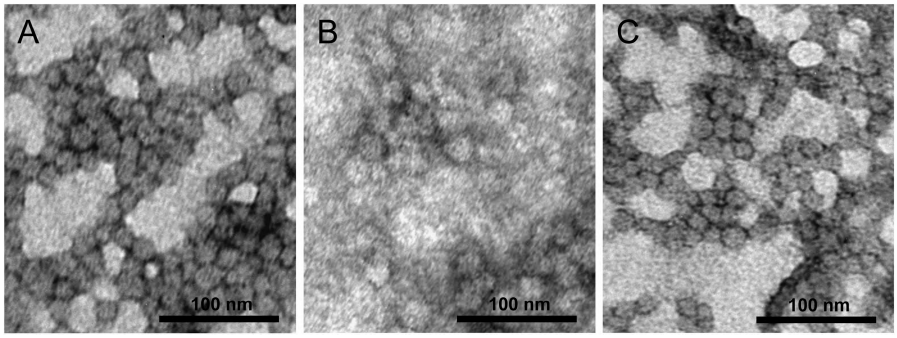

Figure 8 *TEM micrographs of the uranyl acetate stained PFS loaded capsids during the CV experiment. Grids were prepared by quickly dipping the grid into the solution containing the PFS polyanions A) PFS-loaded capsids before the oxidation and reduction cycle B) PFS-loaded capsids after the first oxidation step. C) PFS loaded capsids after the first oxidation and reduction steps.*

Conclusions

The experiments described above have shown that negatively charged PFS is also able to assemble capsid proteins to form T=1 particles; similar to e.g. PSS.[11] As mentioned, this is probably due to the interaction of the negatively charged polymer with the positively charged capsid interior. The redox properties of the encapsulated PFS were quite different from those of the 'free' PFS. The initial shift of the oxidation and reduction potentials might be explained by the interaction of the polymer with the protein shell. In subsequent electrochemical cycles the peaks were found to shift and their intensity was decreased. Apparently, the complex formed between PFS and the capsid protein can be oxidized but not be reversibly reduced. The reduction therefore proceeds incompletely, so that the polymer/protein complex is likely still largely hydrophobic in nature and inaccessible for the subsequent oxidation, also because a substantial part of the polymer is still oxidized. The oxidation peak is thus decreased in height and shifted to a higher voltage because compared to the blank specimen, i.e. 'free' PFS, it is getting more difficult to oxidize the polymer. This cycle repeats itself until all polymer is oxidized and becomes inaccessible for further oxidation. This might explain the CV results, but it still does not explain why the capsids appear to stay intact during this process. It might be that although the PFS is crucial for the initial formation of the T = 1

particle, it is no longer needed when the capsid is formed. Possibly the interactions between the capsid proteins in the formed spherical structure are sufficiently strong to maintain the capsid structure.

Future prospects

Knowledge of the role of negatively charged polyelectrolytes in the stabilization of the capsid after induction of capsid formation might provide novel insight into the interactions between capsid proteins, and thus contribute to the understanding of their assembly mechanisms.

PFS does not seem to be a suitable polyelectrolyte to conduct such studies as the electrochemical oxidation needed to diminish the interactions between PFS and the capsid proteins is experimentally challenging. Therefore a different polymer might be a better candidate for such investigations.

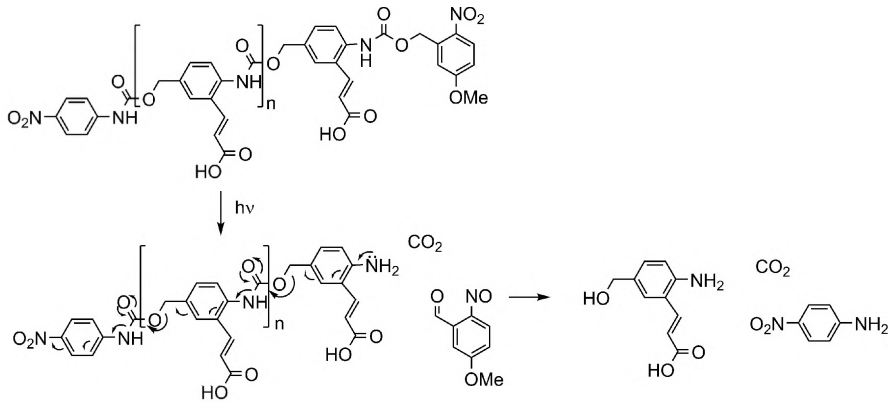

Figure 9 *Self-immolative polymer with photocleavable linker. After UV-irradiation, the photocleavable head-group is cleaved, resulting in a cascade reaction which disassembled the polymer in monomers.*

Sagi et al. [20] developed a self-immolative polymer, which already possesses many of the properties such a polymer should have and which can be modified to fit our requirements. These self-immolative polymers have the same basic features as PFS and PSS; they contain negatively charged carboxylic acids as side chains, which make the polymer water soluble, they could potentially interact with the interior of the capsid, and the polymer can also be synthesized with sulphonated side groups. But more importantly, after removal of the polymer head group, a

chain reaction is set in motion, resulting in the total disassembly of the polymer into its monomers. This would in principle eliminate the stabilizing interactions with the capsid proteins. To make matters even better, the disassembly of the polymer can easily be monitored, as the monomers have a very strong fluorescence emission, and the polymers have not. *Sagi et al.* have also shown that it is possible to add a tail reporter group to the polymer, so that it is possible to visualize the degree of disassembly. The head group triggering the disassembly can easily be varied, and for our purposes a photocleavable head group would be very suitable, since UV-irradiation would probably not be harmful to the capsid proteins (figure 9).

After encapsulation of the immolative polymer, UV-irradiation could release the head group, initiating the disassembly of the polymer. This would then either result in disassembly of the capsid or in an intact, but empty T=1 particle (figure 10).

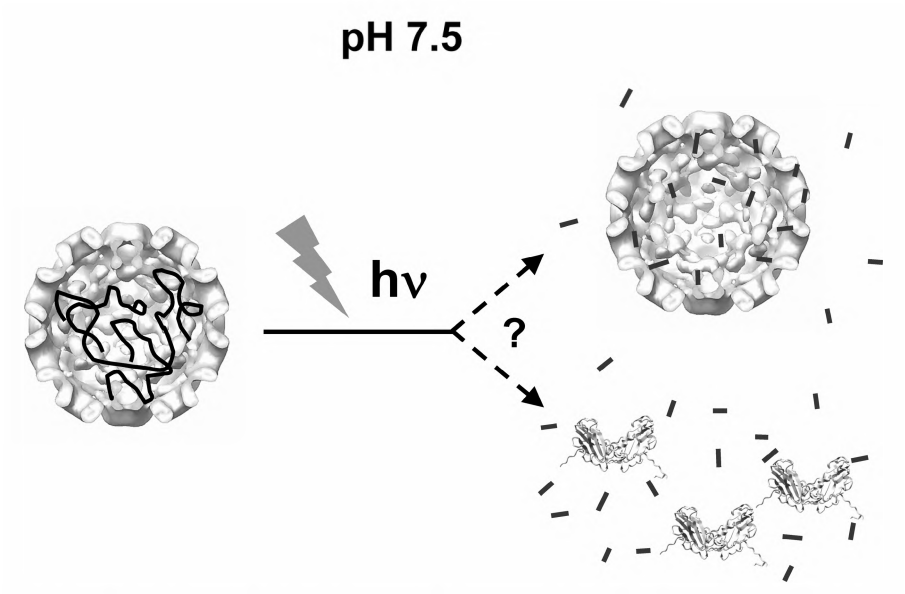

Figure 10 *Schematic representation of the UV-initiated disassembly of the self-immolative polymer shown in figure 9.*

Experimental section

Methods and materials

For all experiments a buffer solution containing 0.4 M NaCl, 0.05 M Tris-HCl, 0.01 M $MgCl_2$, and 0.001 M EDTA of pH 7.5 was used.

Protein concentrations were determined using a Cary 50 Conc (Varian, Middelburg) UV-VIS spectrophotometer. FPLC measurements were performed using a superose 6 PC 3.2/30 analytical column from GE lifesciences, on an Amersham Ettan LC system, fitted with a fractionating device. Buffers for FPLC were filtered with a Millipore 0.2 μM filter before use. TEM grids (Formvar-Carbon) were exposed to an electron discharge treatment using a Cressington Carbon coater and power unit. Unless stated otherwise, the sample was applied to the grids by adding a drop of 5 μL solution (~ 0.2 mg/mL) to the grid and carefully removing it following 1 minute immersion using a filter paper. The grid was allowed to dry for at least 15 minutes before applying 5 μL of a 2% (w/v) uranyl acetate aqueous solution, which was removed after 15 seconds. The grid was again allowed to dry for at least 15 minutes. Samples were studied on a JEOL JEM-1010 TEM (Jeol, Japan). CV measurements were performed on a Autolab PGSTAT10 potentiostat from Echochemie. PFS loaded capsids were purified on FPLC and concentrated to a concentration of ~ 1mg/mL protein with Millipore centrifugal devices (MWCO 100 kDa).

Calculation of PFS encapsulation

The extinction coefficient, ε, of the PFS anions at λ=280 nm was calculated by measuring the UV-Vis absorbance of several samples of PFS polyanions of known concentrations dissolved in MQ. Since the law of lambert-beer states: $A=\varepsilon cd$, where A stands for absorption, c for concentration and d for the pathlenght of the cuvet. The extinction coeffient could thus be calculated.

$$A_{280(PFS)} = \varepsilon_{PFS} c_{PFS} d$$

$$A_{280(CP)} = \varepsilon_{CP} c_{CP} d$$

$$\varepsilon_{PFS} = 2189$$

$$\varepsilon_{CP} = 24075$$

$$\frac{A_{280(PFS)}}{A_{450(PFS)}} = 11.67$$

$$\frac{A_{280(PFS)} + A_{280(CP)}}{A_{450(PFS)}} = 20$$

$$\frac{A_{280(PFS)}}{A_{450(PFS)}} + \frac{A_{280(CP)}}{A_{450(PFS)}} = 20$$

$$\frac{A_{280(CP)}}{A_{450(PFS)}} = 8.33$$

$$A_{280(CP)} = 8.33 . A_{450(PFS)}$$

$$A_{280(CP)} = 8.33 \frac{A_{280(PFS)}}{11.67}$$

$$1.4 A_{280(CP)} = A_{280(PFS)}$$

$$1.4 \varepsilon_{CP} c_{CP} d = \varepsilon_{PFS} c_{PFS} d$$

$$15.4 c_{CP} = c_{PFS}$$

There are 15.4 monomeric PFS units per capsid protein and 924 per T=1 particle. 7200 monomeric units were initially added, hence 12.8% of the added PFS is encapsulated.

The concentration of PFS inside the T=1 capsids is approximately 330 times higher than the expected concentration when the PFS polyanions would randomly be distributed over the entire solution (C_{PFS}). The concentration of PFS inside the T=1 particles ($C_{PFS\_T=1}$) is the number of moles of PFS inside the T=1 particles ($n_{PFS\_T=1}$) divided by the total interior volume of all T=1 particles ($V_{PFS\_T=1}$). Here $n_{PFS\_T=1}$ is the number of moles of PFS added times the percentage that is encapsulated (12.8%). The total interior volume can be calculated by multiplying the number of T=1 particles in solution ($\#_{T=1}$), with the approximated interior volume of one T=1 particle ($V_{T=1}$). The interior volume is approximated by regarding the interior as a perfect sphere and assuming that the capsid wall thickness is the same as it is for the T=3 particles i.e. 10 nm, which leads to a capsid interior radius of 4 nm.

$$V_{T=1} = \frac{4\pi r^3}{3} = 2.68.10^{-22} L$$

$$\#_{T=1} = \frac{C_{CP}}{60} N_A = 1.44.10^{18}$$

$$V_{PFS\_T=1} = V_{T=1}.\#_{T=1} = 3.87.10^{-4} L$$

$$n_{PFS\_T=1} = n_{PFS}.\%_{encapsulated} = 7.39.10^{-6} mol$$

$$C_{PFS\_T=1} = \frac{n_{PFS\_T=1}}{V_{PFS\_T=1}} = 1.91.10^{-2} M$$

$$C_{PFS} = 5.76.10^{-5}$$

$$\text{Concentration factor} = \frac{C_{PFS\_T=1}}{C_{PFS}} = 331$$

References

1. van Delden, R.A., et al., *Unidirectional molecular motor on a gold surface.* Nature, **2005**, *437*, 1337.
2. Vinogradova, O.I., et al., *Multilayer DNA/poly(allylamine hydrochloride) microcapsules: Assembly and mechanical properties.* Biomacromolecules, **2005**, *6*, 1495.
3. Gong, H.F., et al., *Interaction and adhesion properties of polyelectrolyte multilayers.* Langmuir, **2005**, *21*, 7545.
4. Ma, Y.J., et al., *Layer-by-layer constructed macroporous architectures.* Angew. Chem. Int. Edit., **2007**, *46*, 1702.
5. Johnston, A.P.R., E.S. Read, and F. Caruso, *DNA multilayer films on planar and colloidal supports: Sequential assembly of like-charged polyelectrolytes.* Nano Lett., **2005**, *5*, 953.
6. Johnston, A.P.R., A.N. Zelikin, and F. Caruso, *Assembling DNA into advanced materials: From nanostructured films to Biosensing and delivery systems.* Adv. Mater., **2007**, *19*, 3727.
7. Vriezema, D.M., et al., *Self-assembled nanoreactors.* Chem. Rev., **2005**, *105*, 1445.
8. Kanaras, A.G., et al., *Towards multistep nanostructure synthesis: Programmed enzymatic self-assembly of DNA/gold systems.* Angew. Chem. Int. Edit., **2003**, *42*, 191.
9. Xu, X.Y., et al., *Asymmetric functionalization of gold nanoparticles with oligonucleotides.* J. Am. Chem. Soc., **2006**, *128*, 9286.
10. Liepold, L.O., et al., *Structural transitions in Cowpea chlorotic mottle virus (CCMV).* Physical Biology, **2005**, *2*, S166.

11. Sikkema, F.D., et al., *Monodisperse polymer-virus hybrid nanoparticles.* Org. Biomol. Chem., **2007**, *5*, 54.
12. Zhong, Z.Y., et al., *Water-soluble cationic poly(ferrocenylsilane): An efficient DNA condensation and transfection agent.* Journal of Controlled Release, **2006**, *116*, e81.
13. Whittell, G.R. and I. Manners, *Metallopolymers: New multifunctional materials.* Adv. Mater., **2007**, *19*, 3439.
14. Ma, Y.J., et al., *Redox-controlled molecular permeability of composite-wall microcapsules.* Nat. Mater., **2006**, *5*, 724.
15. Ma, Y.J., M.A. Hempenius, and G.J. Vancso, *Electrostatic assembly with poly(ferrocenylsilanes).* J. Inorg. Organomet. Polym. Mater., **2007**, *17*, 3.
16. Meyer, T.E. and M.D. Kamen, *New Perspectives on C-Type Cytochromes.* Adv.Protein Chem., **1982**, *35*, 105.
17. Verduin, B.J.M., *Preparation of Ccmv-Protein in Connection with Its Association into a Spherical-Particle.* FEBS Lett., **1974**, *45*, 50.
18. Verduin, B.J.M., *Degradation of Cowpea Chlorotic Mottle Virus Ribonucleic-Acid Insitu.* J. Gen. Virol., **1978**, *39*, 131.
19. Manners, I., *Ring-opening polymerization of metallocenophanes: A new route to transition metal-based polymers*, in *Advances in Organometallic Chemistry, Vol 37*. 1995, Academic Press Inc: San Diego. p. 131.
20. Sagi, A., et al., *Self-immolative polymers.* J. Am. Chem. Soc., **2008**, *130*, 5434.

Chapter 4

Virus-like particles templated by DNA micelles *

Introduction

In the previous chapter it was shown that different kinds of negatively charged polyelectrolytes, i.e. PSS and PFS, can be loaded into the CCMV capsid via a facile route. The negative charges presumably interact with the positively charged capsid interior, which prevents the polyelectrolyte from diffusing out of the capsid through the pores. They possibly also provide stabilizing interactions to keep the capsid intact. Not all compounds possess these properties, which makes their encapsulation more challenging. Drugs are typical examples of compounds that possess none of the above mentioned properties. The development of a method to encapsulate these small compounds could, however, be of great importance for the development of a capsid-based drug delivery system. The use of plant viruses as drug-delivery vehicles has several advantages. Plant viruses are in the nano-meter size range, which enhances their cell-permeability.[1-3] Furthermore they are very monodisperse, biodegradable and biocompatible. Their exterior can be modified to contain for instance cell-targeting ligands, or polyethylene glycol (PEG) chains to suppress the immune response[4] and to prolong blood circulation times. The porous nature of CCMV however, makes it hard to retain small drug-like molecules inside the capsid without for instance covalently attaching them to the capsid proteins. This is difficult because, drug molecules are often hydrophobic and therefore solvent incompatibilities severely complicate covalent attachment to the capsid interior.

* This work was carried out in collaboration with Minseok Kwak and Andreas Hermann from the University of Groningen and Melanie Brasch from the University of Twente, and has been published: *JACS*, **2010**, 132, 7834

It is however possible to entrap hydrophobic molecules in micelles consisting of DNA amphiphiles. These amphiphiles consist of single stranded DNA coupled to either fatty chains or hydrophobic organic polymers. They aggregate into micellar structures with a hydrophobic interior and an anionic corona. The micelles are approximately 10 nm in diameter, which roughly corresponds to the internal diameter of a CCMV capsid. It is therefore conceivable that the anionic corona can interact with the capsid proteins and induce assembly. While hydrophobic molecules can be entrapped in the hydrophobic interior of the micelles, small hydrophilic molecules can be attached to the micelles by linking them to a single stranded DNA chain, which is subsequently hybridized with the micellar DNA.
In this chapter a strategy along this line, i.e. the facile self-assembly and loading of CCMV capsids with small molecules using DNA amphiphiles is reported. (figure 1)

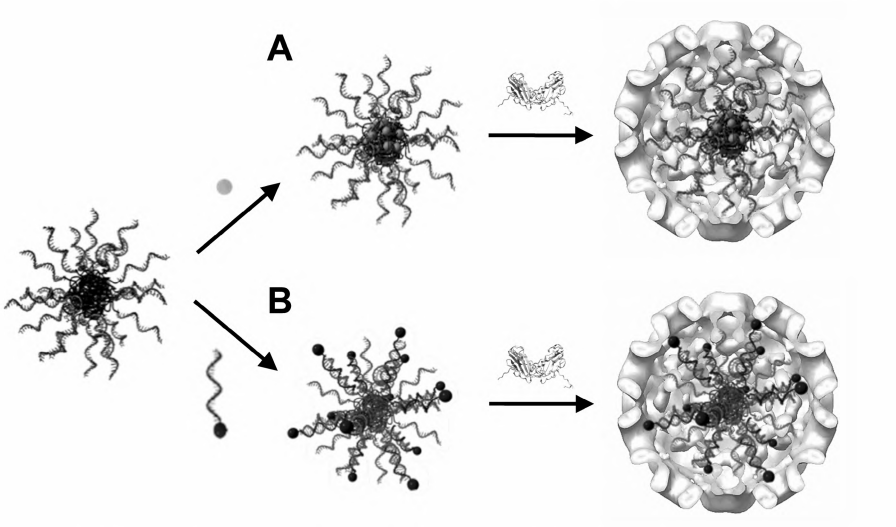

Figure 1 *DNA micelle-templated capsid formation. (A) Loading of hydrophobic molecules into the micellar core followed by encapsulation. (B) Equipping functional DNA attached moieties to the micelle via hybridization followed by encapsulation.*

Results and Discussion

Two different classes of DNA amphiphiles were used for this research. The first class consists of low molecular weight hydrophobic molecules that are attached to small oligonucleotides (ODNs).[5] A lipid-DNA 11-mer (UU11) containing two 5-dodec-1-ynyluracil nucleobases at the 5'- end followed by 9 naturally occuring nucleobases was synthesized and used for this purpose (figure 2).

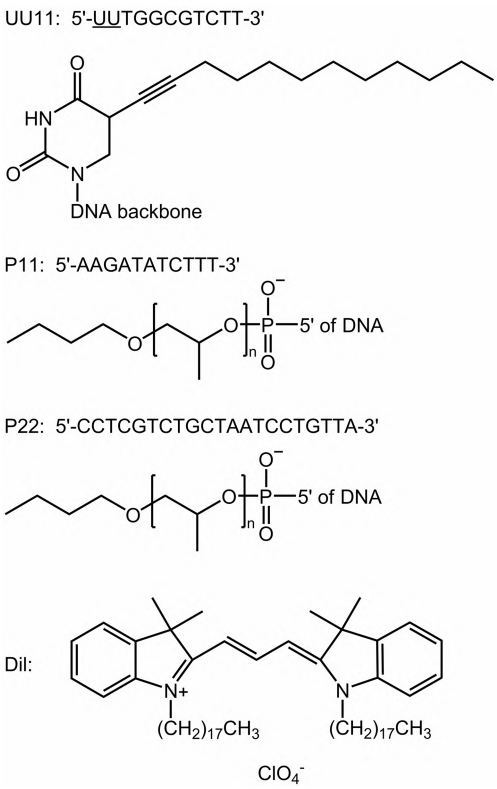

Figure 2 *DNA amphiphiles used for encapsulation and structure of DiI dye.*

The second class consists of DNA block copolymers (DBCs), in which a nucleic acid sequence is covalently attached to a hydrophobic organic polymer via its end group.[6] Here two different DBCs were used; both containing polypropylene oxide (PPO) blocks with a molecular weight of 6800 g/mol, but attached to DNA of different length, i.e. 11 nucleotides (P11) and 22 nucleotides (P22) (figure 2). All three amphiphiles form micelles with a diameter of 7-11 nm at room temperature.

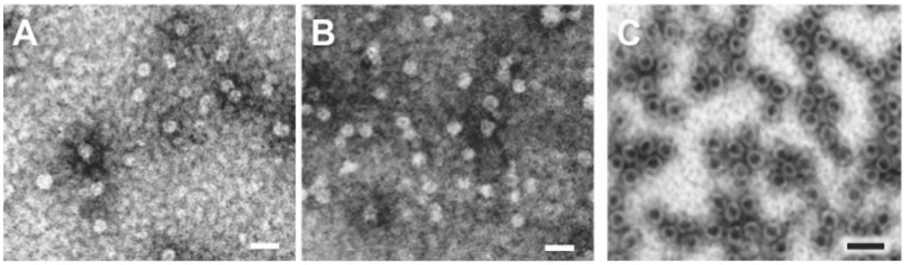

Figure 3 *TEM micrographs of DNA-micelle-capsid assemblies. The scale bars are 40 nm. A) P11-capsid protein B) P22-capsid protein C) Wild type empty capsid assembled at pH 5.0. Staining is by uranyl acetate.*

To test whether the micelles could induce capsid formation, a similar procedure was followed as for capsid formation using PFS or PSS. The capsid proteins were mixed with the DNA amphiphiles at pH 7.5 in various ratios, and the results were analyzed by FPLC or TEM (figure 3). A ratio of ~2.3 amphiphiles to one capsid protein monomer was found to be the optimal ratio for induction of capsid formation.

To distinguish between capsid protein and micelles on FPLC and to demonstrate that it was possible to incorporate small hydrophobic molecules inside the capsid, the micelles were loaded with a test compound, i.e. a hydrophobic dye, 1,1'-dioctadecyl-3,3,3',3'-tetramethylindocarbocyanine perchlorate (DiI) (figure 2), before mixing with the capsid proteins and inducing capsid formation. Non-encapsulated micelles eluted at 1.6 mL (figure 4), while the sample containing both capsid protein and DiI-loaded UU11 micelles showed a peak at V=1.8 mL, which is the unassembled capsid protein, and an extra peak at V=1.4 mL. The latter peak also showed DiI-specific absorption at λ=549 nm, and an increased absorption at λ=280 nm with respect to the absorption at λ=549 nm, indicating the presence of protein, which absorbs strongly at λ=280 nm. As the capsid protein is not able to assemble at pH 7.5, the micelle must have induced the assembly of the capsid. Together with the observed co-elution of capsid protein and DiI-loaded micelle, this strongly suggests that the micelle is encapsulated inside the capsid. As the formed particles elute at V=1.4 mL, they are probably smaller than the native virus, which elutes at V=1.1 mL.

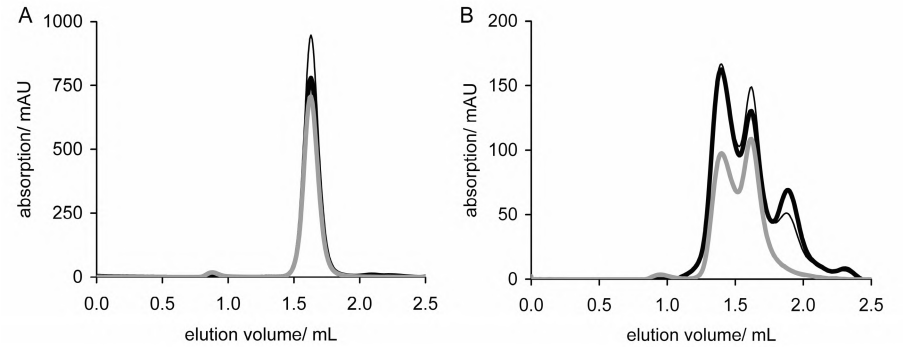

Figure 4 *FPLC traces of micelles and micelle-capsid assemblies. The FPLC system was equipped with a Superose 6 column and equilibrated with a buffer of pH 7.5. The thin line represents UV-Vis absorption at λ=260 nm, the thick black line UV-Vis absorption at λ=280 and the thick grey line represents DiI specific absorption at λ=549nm. A) DiI-loaded UU11 micelles. B) Capsid proteins incorporating DiI loaded UU11 micelles.*

To further determine the size of the micelle-capsid assembly, the FPLC fractions were analyzed by TEM.

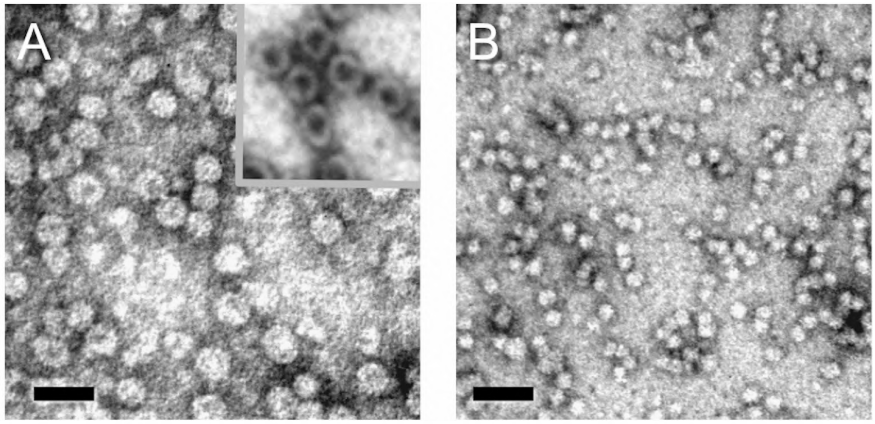

Figure 5 *TEM micrographs stained by uranyl acetate of UU11-capsid FPLC fractions. Scale bars are 40 nm. A) Capsid-UU11 micelle particles eluting at V=1.4 mL. The inset shows empty capsids for comparison. B) UU11 micelles eluting at V=1.6 mL.*

This revealed that the fraction eluting at V=1.4 formed structures of 19.9 ±3.1 nm (figure 5), which is slightly larger than would be expected for a T=1 particle. It is possible, therefore, that the capsid proteins (partially) adopt a T=2 symmetry. The interior of the capsids appeared to be unstained, indicating they have a filled interior.

The non-encapsulated micelles eluting at V=1.6 were 8.1±1.6 nm in size, which corresponds with their known size.[7] The diameter of the micelles is 11 nm smaller than the diameter of the formed capsids. Together with the thickness of the capsid wall, which is approximately 5 nm, this leads to a good fit of the micelle inside the capsid particle.

To confirm the presence of capsid protein in the FPLC fraction eluting at V=1.4 mL, the fractions were analyzed by SDS-Page and visualized using silver staining (figure 6). This showed the presence of capsid protein in the V=1.4 mL fraction, (lane 3).

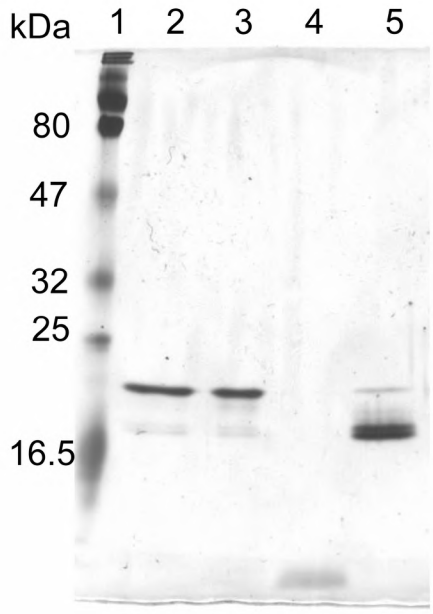

Figure 6 *Silver stained SDS-Page gel of FPLC fractions. Lane 1: marker, lane 2: wt CP, lane 3: capsid-UU11 micelle particles eluting at V=1.4 mL, lane 4: DiI loaded UU11 micelles eluting at V=1.6 mL.*

To further explore the use of the capsid-micelle particles as vehicles that can load small compounds, UU11 micelles were loaded with pyrene, a hydrophobic small fluorescent dye. Particles were generated using the same procedure as for the DiI-loaded capsids. FPLC analysis showed a clear pyrene absorption at V=1.3 mL (figure 7a). The fluorescent spectrum of this fraction showed characteristic pyrene emission bands (figure 7b).

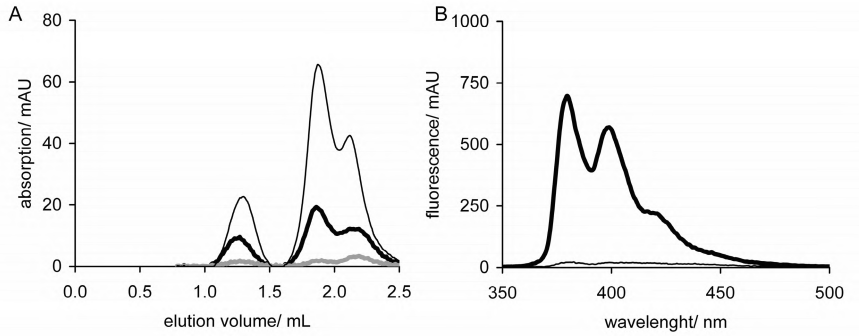

***Figure** 7 FPLC trace and fluorescence spectra of pyrene-loaded capsid-UU11 micelle particles. A) FPLC trace. The thick black line represents protein absorption at λ=280 nm, the grey line represents pyrene absorption at λ=342 nm, the thin black line represents absorption of the Dylight™ labeled wt CP at λ=654 nm. B) Fluorescence spectra (λ_{ex} = 342 nm) of the FPLC fraction eluting at V=1.3 mL as depicted in figure 7A (thick black line), and wt CP (thin black line).*

In order to investigate the loading of the micelle-capsid particles with small hydrophilic molecules, UU11 micelles were equipped with 6-carboxylic-X-rhodamine (ROX), a small hydrophilic dye. This was achieved by hybridizing a single DNA strand conjugated to ROX with the UU11 micelles (figure 1b). The capsid proteins were labeled with a Dylight™ 647 dye prior to mixing with the labeled micelles. FPLC analysis confirmed the presence of ROX in the formed particles (figure 8).

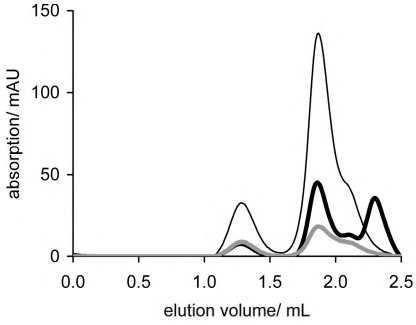

***Figure 8** FPLC trace of ROX loaded UU11 micelle-capsid particles. The FPLC system was equipped with a Superose 6 column and equilibrated with a buffer of pH 7.5. The thick black line represents the absorption at λ=280, the thin black line represents absorption of the Dylight™ labeled wt CP at λ=654 nm. and the grey line represents ROX specific absorption at λ=575nm.*

Using the absorption values obtained by FPLC, the aggregation number of the UU11 micelle was calculated to be 25 ±2. Fifty percent of the DNA was hybridized to the ROX containing complementary DNA sequence, and thus has a double charge density, so the total amount of negative charges per capsid is approximately 412. Assuming a T=2 conformation, there would be 3.4 negative charges per capsid protein. This is rather similar to the number of negative charges found for the PFS containing capsids, which is 5.1 (see experimental section for calculation details).

Conclusions

The incorporation of negatively charged micelles inside CCMV-capsids demonstrates that not only linear negatively charged polyelectrolytes are capable of inducing the assembly of the capsid proteins. The results presented in this chapter further demonstrate the possibilities to direct and control the assembly of the CCMV capsid proteins. It is rewarding that through the encapsulation of amphiphilic micelles the co-encapsulation of small molecules can be achieved. The general loading strategy presented in this chapter might be of great value for the development of new capsid-based drug delivery systems.

Experimental section

Methods and materials

For all experiments a buffer of pH 7.5 containing 0.5 M NaCl, 0.05 M Tris-HCl, 0.01 M $MgCl_2$, and 0.001 M EDTA at pH 7.5 was used.

Protein concentrations were determined using a Cary 50 Conc (Varian, Middelburg) UV-VIS spectrophotometer.

FPLC measurements were performed using a Superose 6 PC 3.2/30 analytical column from GE lifesciences, attached to an Amersham Ettan LC system, fitted with a fractionating device. Buffers for FPLC were filtered with a Millipore 0.2 µM filter before use. TEM grids (Formvar-Carbon) were exposed to an electron discharge treatment using a Cressington Carbon coater and power unit. Unless stated otherwise, the samples were applied to the grids by adding a drop of 5 µL solution (~ 0.2 mg/mL) to the grid and carefully removing it using a filter paper after 1 minute immersion. The grid was allowed to dry for at least 15 minutes before applying 5 µL of a 2% (w/v) uranyl acetate aqueous solution, which was removed after 15 seconds. The grid was again allowed to dry for at least 15 minutes. Samples were studied on a JEOL JEM-1010 TEM (Jeol, Japan) instrument. Fluorescence measurements were performed on a Perkin Elmer LS-55 fluorescence spectrometer.

DNA amphiphiles

DNA amphiphiles were synthesized and analyzed by Minseok Kwak as described in the literature.[7,8]

Production and purification of wild type capsid protein (wt CP)

The purification of the CCMV virus and the removal of its RNA were carried out according to literature procedures.[1]

Wt CP labeling

A solution of capsid protein in 0.1 M phosphate buffer pH 7.5, 0.15 M NaCl, was added to a vial containing 50 µg DyLight™ 647 NHS-ester and stirred for 1 h at room temperature (RT). The non-coupled dye was removed by dialysis to a buffer pH 7.5 or pH 5.0.

Preparation of micelle-loaded capsids

A DNA amphiphile and wt CP were mixed in a 1 : 2.3 molar ratio (Table 1) in pH 7.5 capsid buffer and the mixture was allowed to form capsids for 1 h at 4 °C.

Table 1 *Mixed volume and concentration of micelle loaded capsids*

Amphiphile used for micelle loaded capsid formation (molar ratio of the amphiphile : wt CP)	DNA amphiphile	Wt CP
Pyrene/UU11 (1 : 2.3)	19.6 µL (209 µM)	100 µL (94µM)
DiI/UU11 (1 :3.6)	8.1 µL (400 µM)	100 µL (148 µM)
ROX-ODN/UU11 (1 : 2.3)	19.6 µL (209µM)	90 µL (83 µM)

Preparation of Loaded Micelles

DiI or pyrene was dissolved in acetone. A drop of the dye solution (400 µg/ml) was added to a tube and subsequently evaporated. A UU11 micellar solution was added to the tube and the mixture was agitated overnight at room temperature.

UU11 was hybridized with 50 mol percent of ROX-ODN in $MgAc_2$ (10 mM) buffer at pH 7.5.

Calculation of the aggregation number (Z) of the UU11 micelle inside the capsid

The molarity ratio of UU11 : wt CP was calculated using the Beer-Lambert law, $A = \varepsilon cl$. From the absorbances of ROX on complementary DNA (4.30 mAU, 50% hybridized) and Dylight^{TM} 647 on wt CP (85.5 mAU, 17% labeled), see elution peak at 1.28 ml in Figure 8, a molar ratio of 1 : 4.82 was calculated. In a capsid of T = 2 geometry, which consists of 120 CPs, this corresponds to 120/4.82=24.9 single-stranded UU11 molecules in the capsid-cavity. The micelle in the cavity thus has an aggregation number of Z = 24.9.

Calculation of the number of negative charges per capsid protein for encapsulated PFS

As stated in the previous chapter, the concentration of PFS inside the capsid is 1.9×10^{-2} M. The number of molecules of PFS molecules per capsid can be obtained from this number, by multiplying with the internal volume of a T=1 capsid (assuming a perfect sphere with a diameter of 8 nm), and with the number of Avogadro. This leads to 3.08 molecules of PFS/capsid. Since PFS has an average charge density of 100 negative charges per molecule, the number of charges per 60 capsid proteins (T=1 particle) is 308, and thus the number of charges per capsid protein is 5.1.

References

1. Allen, T.M. and P.R. Cullis, *Drug delivery systems: Entering the mainstream.* Science, **2004**, *303*, 1818.
2. Holder, P.G. and M.B. Francis, *Integration of a self-assembling protein scaffold with water-soluble single-walled carbon nanotubes.* Angew. Chem. Int. Edit., **2007**, *46*, 4370.
3. Maeda, H., et al., *Tumor vascular permeability and the EPR effect in macromolecular therapeutics: a review*. Journal of Controlled Release, **2000**, *65*, 271.
4. Kovacs, E.W., et al., *Dual-surface-modified bacteriophage MS2 as an ideal scaffold for a viral capsid-based drug delivery system.* Bioconjugate Chem., **2007**, *18*, 1140.
5. Xu, C., et al., *Microcontact printing of DNA-surfactant arrays on solid substrates.* Journal of Materials Chemistry, **2003**, *13*, 3044.
6. Alemdaroglu, F.E., et al., *DNA-Templated Synthesis in Three Dimensions: Introducing a Micellar Scaffold for Organic Reactions.* Angew. Chem. Int. Edit., **2006**, *45*, 4206.
7. Kwak, M., et al., *Virus-like Particles Templated by DNA Micelles: A General Method for Loading Virus Nanocarriers.* J. Am. Chem. Soc., **2010**, *132*, 7834.
8. Alemdaroglu, F.E. and A. Herrmann, *DNA meets synthetic polymers - highly versatile hybrid materials.* Org. Biomol. Chem., **2007**, *5*, 1311.

Chapter 5

*Bacterial capsid protein**

Purification and assembly of the His-tag capsid protein

The CCMV capsid protein dimers can form a plethora of different structures when the ionic strength, the concentration of divalent cations and the pH of the buffer solution are varied.[1] This demonstrates the enormous versatility of the CCMV capsid protein to self-assemble. To further explore and expand the (assembly) properties of the capsid protein, it was heterologeously expressed in *E.coli* to be able to genetically modify the capsid protein. The procedure involved the extraction of RNA from the CCMV virus particles after which a reverse transcriptase reaction was used to obtain the copy DNA (cDNA). The cDNA was amplified using PCR and cloned into a pET-15b *E.coli* expression vector. The obtained plasmid was transformed into *E.coli* BL21(DE3)pLysS cells. This work will be described in detail in the forthcoming thesis of L.J.A. Hendriks.[2]

Importantly, the expression vector contained a histidine-tag (His-tag) sequence which allows for easy purification of the protein, followed by a thrombin cleavage sequence, which facilitates removal of the His-tag. The DNA encoding for the coat protein was inserted directly after the thrombin cleavage site, and thus the resulting coat protein contained an N-terminal His-tag followed by a thrombin cleavage site.

A His-tag is an amino acid motif which consists of several histidines in a row, in the present case 6. It is widely used for affinity purification of proteins. Histidines are good ligands to nickel, and in this way the engineered proteins can be selectively bound to a solid support functionalized with nickel chelating nitrilotriacetic acid (NTA) groups, i.e. Ni-NTA resins.

* Part of this work is submitted for publication

Protein expression and purification

For capsid protein with His-tag (His-CP) expression, *E.coli* cells suitable for high-level protein expression (BL21 cells) containing the coat protein plasmid were grown overnight in lysogeny broth (LB) medium, a standard bacterial growth medium, provided with the appropriate antibiotics. The next day the culture was diluted 10 times with LB medium and antibiotics. Expression of the protein was induced by adding isopropyl β-D-1-thiogalactopyranoside (IPTG) at the logarithmic growth phase of the cells and after approximately 7 hrs the cells were pelleted by centrifugation and stored at -20°C awaiting purification. The production of His-CP increased more than six fold when cells were grown at 30°C instead of 37°C, resulting in final yields of approximately 20 mg of His-CP from a 1 L bacterial culture.

For His-CP purification, the pelleted cells were defrosted on ice, and resuspended in a lysis buffer containing a small amount of imidazole. Imidazole has the same structure as histidine side groups, and is therefore able to coordinate to the Ni-NTA resin. The addition of imidazole to the lysis buffer prevents some a-specific binding of proteins to the Ni-NTA resin, while the concentration is still sufficiently low not to interfere with the binding of the engineered protein to the resin. To disrupt the bacterial cell walls, lysozyme was added, followed by sonication. The resulting homogeneous mixture was subjected to centrifugation, to pellet the bacterial debris. The supernatant containing soluble proteins was added to the Ni-NTA column and the proteins were allowed to bind for at least an hour at 4°C. The flow through was then collected, and the column was washed with a wash buffer containing a larger concentration of imidazole than the lysis buffer, to remove all non-specifically bound proteins from the resin. To elute His-CP from the column, small amounts of elution buffer containing a large excess of imidazole were added. The imidazole competes with the His-tagged proteins for binding to the column, and since a large excess is used, the proteins are eluted this way. The elution fractions were then dialyzed to buffer pH 7.5, to remove excess of imidazole in the buffer. This step should take place as soon as possible, but at least within a day after the purification, since high concentrations of imidazole are harmful to proteins.

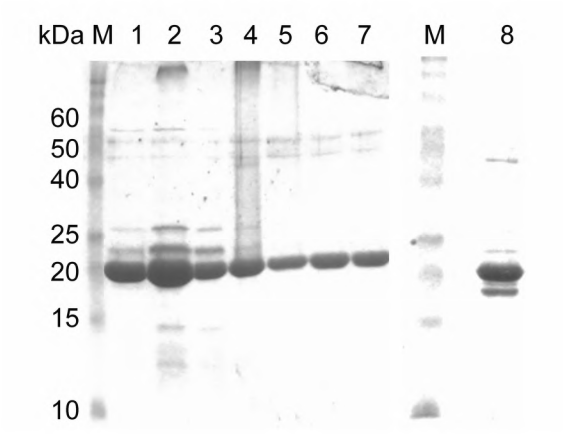

Figure 1 *SDS-PAGE analysis of purified His-CP samples. M=Marker. Lane 1-7 are consequtive elution fractions of a His-CP containing protein mixture, after His-CP elution from a Ni-NTA column. Lane 8 is the same sample as 1-7, after 46 days. A small band which is degraded His-CP protein, has appeared underneath the main His-CP band.*

After purification a portion of the His-CP protein appeared to precipitate within a few days. This seemed to be somewhat concentration dependent. The soluble part of the solution did not appear to contain much degraded protein as was shown by SDS-PAGE analysis (figure 1). It was tried therefore to assemble the soluble part of the His-CP solution into capsid particles. Unfortunately, the His-CP proteins rarely assembled into capsids, as was evident from the absence of capsid particles on the TEM pictures and the absence of a capsid peak at V=1.1 mL in the FPLC traces. Incidentally FPLC traces showed traces of assembly, but the peak height at V=1.1 mL was much smaller than the peak height of a sample of wild-type (wt) CP with a similar concentration. The inability to assemble was not caused by degradation of the proteins, since SDS-PAGE analysis showed no degradation and only proteins of the correct molecular weight were observed (figure 1). Comparison of the UV-Vis spectra of the wild-type capsid protein (wt CP) and His-CP showed that His-CP had a much stronger absorption at λ = 260 nm compared to the wt CP. Nucleic acids absorb mainly at λ = 260 nm and proteins at λ = 280 nm, and therefore the ratio of these absorption values (the 260/280 ratio) is indicative of the purity of the protein. The wt CP has a 260/280 ratio of 0.7, but the ratio for His-CP was usually between 1.1 and 1.6.

It is very likely, therefore, that the His-CP samples contained more nucleic acid material* than the wt CP samples. Probably, the nucleic acid material originated from *E.coli*, and co-eluted from the Ni-NTA column, due to electrostatic interactions between the positively charged N-terminus of His-CP and the negatively charged nucleic acid backbone. To examine if this was the reason for the poor assembly of His-CP, purification conditions had to be found that would avoid this co-isolation of nucleic acid material with His-CP. Conditions that would destroy the nucleic acid material, i.e. shearing and the addition of nucleases, were complemented with a range of conditions that had the purpose of separating the remains of the nucleic acids from the proteins (table 1). Some of these separating conditions aimed at disrupting the presumed electrostatic interactions between the nucleic acid material and His-CP, either by increasing the ionic strength of the buffer solutions used, or by increasing the pH to deprotonate the positively charged residues of His-CP. Others intended to disturb hydrophobic interactions that might exist between the nucleic acids and His-CP, and also methods to precipitate the nucleic acid material were studied.

Table 1 *Different purification conditions used for the optimization of His-CP purification.*

#	Condition	Purpose
1	Normal, but with extensive resuspending of pelleted cells	Destroying nucleic acid material by shear force
2	As 1 + addition of RNAse and DNAse to the pelleted cells after sonication	Cleaving nucleic acid material
3	As 2 + addition of 0.8% (v/v) Triton X-100 to all buffers	Disrupting hydrophobic interactions
4	As 2 + addition of 2 M NaCl to all buffers instead of 0.3 M normally present	Disrupting ionic interactions
5	As 2 + addition of 10% EtOH (v/v) to all buffers	Disrupting hydrophobic interactions
6	As 2 + addition of 30% Glycerol (v/v) to all buffers	Disrupting hydrophobic interactions
7	As 1 + addition of 0.04% poly(ethyleneimine) at pH 8.0 to pelleted cells for 15' on ice before sonication	Precipitating nucleic acid material with poly(ethyleneimine)
8	As 2 + addition of 2 M $MgCl_2$ to all buffers	Disrupting ionic interactions
9	As 1 + addition of NaOH to increase the pH to 12	Disrupting ionic interactions by deprotonating arginines and lysines

* It should be noted that the co-isolated impurity to the CP might be more complex than just nucleic acid. Throughout this chapter we will refer to this impurity as nucleic acid material

Table 2 *Concentrations and 260/280 ratio of the purified His-CP proteins using the purification conditions mentioned in table 1.*

#	Condition	260/280 Ratio	Concentration sample (mg/mL)	Concentration applied to TEM grid (mg/mL)
1	Extensive resuspending	1.14	1.09	0.03
2	RNAse and DNAse	0.96	0.66	0.02
3	0.8% (v/v) Triton X-100	0.61 [a]	2.26	0.45
4	2 M NaCl	0.72	2.16	0.05
5	10% EtOH (v/v)	1.41	1.31	0.26
6	30% Glycerol (v/v)	1.54	1.72	0.34
7	0.04% PEI	1.56	1.20	0.24
8	2 M $MgCl_2$	0.84	1.12	0.03
9	pH to 12	1.32		

[a] Unreliable UV-Vis values, Triton X-100 was not completely removed from the sample.

To measure which purification conditions were best suited for the removal of nucleic acid material from His-CP, the 260/280 ratio was determined for all samples. As expected, the purification conditions aimed at disturbing the electrostatic interactions were most successful (table 2, entries 4 and 8). The 260/280 ratios of these samples were approximately the same as for the wt CP. On the other hand, the purification conditions that disturbed hydrophobic interactions even seemed to have a disadvantageous effect on the 260/280 ratio (table 2, entries 5 and 6).

To test if the modified purification conditions still resulted in pure and intact protein, the samples were analyzed by SDS-PAGE. This showed that only the last purification condition resulted in degraded protein (figure 2). This was most likely caused by degradation of the protein due to the high pH of the buffer solutions used for these samples. This pH was raised to 12, in order to deprotonate the positively charged residues of His-CP in that way disrupting interactions with the negatively charged nucleic acids.

All the other samples on the gel seemed to consist of intact and pure proteins of the correct molecular weight. The proteins were not dialyzed to a capsid buffer pH 7.5 prior to their application on the gel resulting in some smearing of the lanes, which could be prevented by dialysis to a buffer with lower ionic strength (not shown).

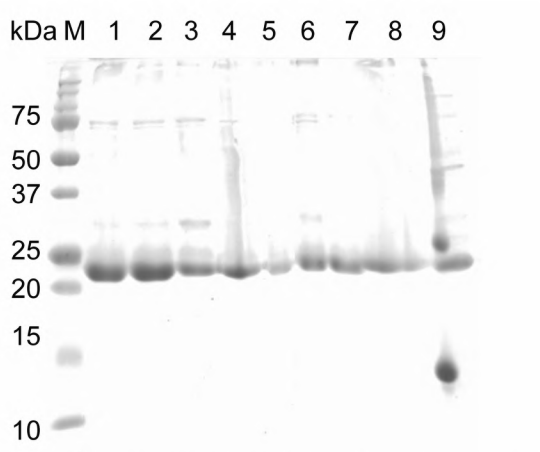

Figure 2 *SDS-PAGE gel of His-CP purified with different purification conditions. M=Marker. Lane 1-9 correspond to purification conditions 1-9 (table 1).*

To test whether the 260/280 ratio indeed had an effect on the capsid assembly, the different samples were dialyzed to pH 5.0 and all samples were analyzed by TEM (figure 3). Samples with a 260/280 ratio < 0.9 not only contained far more capsids than the samples with a 260/280 ratio > 0.9, but the interior of the capsid was also empty, instead of filled.

An empty capsid interior is visible in the TEM analysis as a dark spot inside the bright capsid shell. This effect is caused by the procedure used to provide contrast to the materials in the EM, i.e. negative staining using uranyl acetate. Proteins, in general, are hardly stained, and therefore appear as bright spots on a dark background. When the interior of the capsid is filled, for example with nucleic acids (as is the case for the CCMV virus), the uranyl acetate can not freely enter the capsid interior and the virus is therefore visible as a filled white circle (figure 3a).The staining agent can, however, enter the interior of the empty capsid, which is therefore visible as a dark spot surrounded by a white rim (figure 3b). It seems that the nucleic acid material can accumulate inside the capsids, since samples with a high 260/280 ratio appear to have a filled interior (figure 3 e, g-i). In contrast to the native RNA, this does not seem to drive the assembly of the capsid, since less capsids are formed in samples which have a 260/280 ratio > 0.9 compared to the samples which have a 260/280 ratio < 0.9. Although the amount of capsids formed is difficult to determine using TEM, the concentration of sample that is needed to get TEM images with a well-distributed capsid density does give an indication.

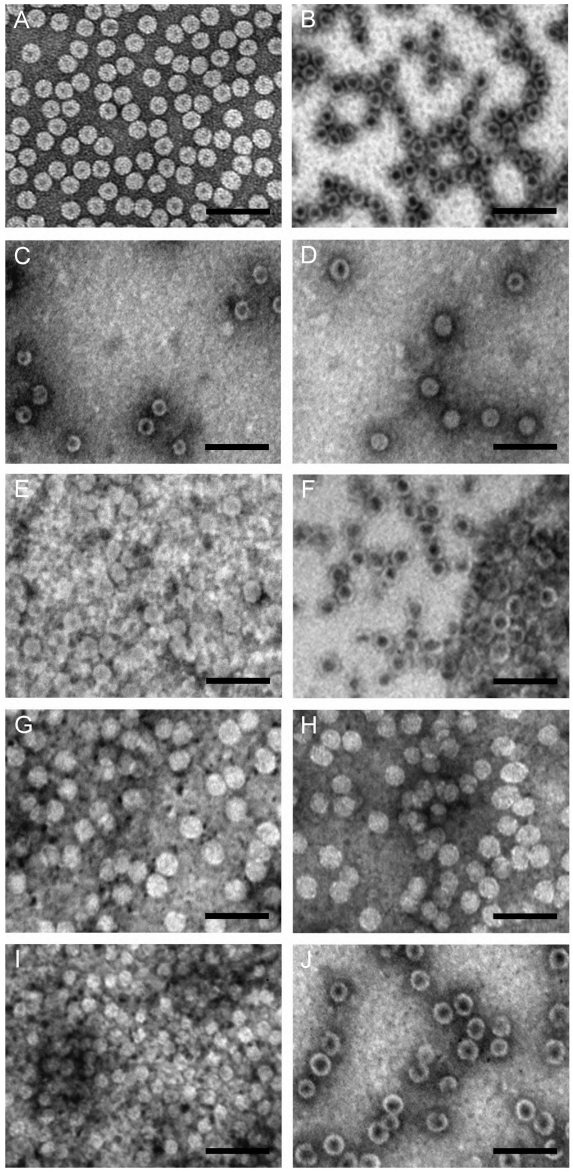

Figure 3 *TEM pictures of assembled His-CP purified following different purification conditions (see table 1). A) CCMV virus. B) wt CP. C) His-CP purified by extensive resuspending. D) His-CP purified by adding nucleases. E) His-CP purified with additional Triton X-100. F) His-CP purified with buffers containing 2 M NaCl. G) His-CP purified with buffers containing 10% EtOH. H) His-CP purified with buffers containing 30% glycerol. I) His-CP purified by addition of poly(ethyleneimine) to pelleted cells. J) His-CP purified with buffers containing 2M $MgCl_2$. Scale bar = 100 nm.*

In order to quantify the amount of capsids that were formed more precisely, some samples were also analyzed by FPLC (figure 4). This analysis was not possible for all samples; since His-CP substantially precipitated from the samples with a 260/280 ratio > 0.9. The FPLC trace of sample 2, which was only treated with nucleases, showed a minor capsid peak at V=1.1 mL and hardly any capsid dimers. Although this might suggest that no protein was present in the sample, UV-Vis measurements showed that the sample concentration was approximately 1 mg/mL. When the assembly efficiency of proteins obtained by purification procedures 2 and 4 (table 1) are compared by FPLC (figure 4) it is clear that procedure 4 gives a much higher yield of capsids than the other procedures. It seems therefore that the purification conditions applied to sample 4 greatly increase the ability of the His-CP proteins to self-assemble. In the experiments so far 2 M of NaCl was used in all buffers. In order to further optimize this procedure, different concentrations of NaCl were tried. All NaCl concentrations between 1 M and 2 M gave the same result. For the remaining His-CP purifications described in this thesis, buffers with 1.5 M NaCl were therefore used.

The increase in ionic strength of the buffers used in the purification procedure seems to be an efficient and simple way to avoid co-purification of nucleic acid material with His-CP. Moreover, it seems that the presence of the non-native nucleic acid material prevented the proper assembly of the capsid, and might also have caused the rapid precipitation of His-CP over time.

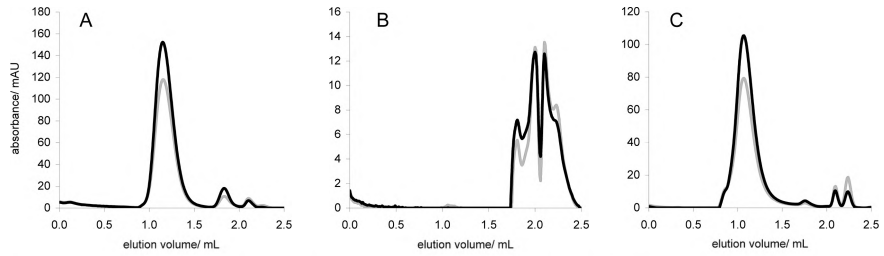

Figure 4 *FPLC graphs of assembled wt CP and His-CP at pH 5.0. Samples were measured using an FPLC equipped with a superose 6 column. Black lines represent absorption at λ=280 nm, grey lines represent absorption at λ=260 nm. A) wt CP. B) His-CP purified with nucleases (table 1, entry 2). C) His-CP purified using buffers with 2 M NaCl added (table 1, entry 4).*

Degradation of capsid proteins

Like almost all proteins, the capsid protein is sensitive to degradation. Especially the N-terminus of the capsid protein is sensitive to this process. At pH 5.0 the degradation is very slow- the capsid is stable for months, while at pH 7.5, the degradation is quite rapid –after 2 days substantial degradation product is observed. The 'protecting' effect of the pH might be due to inactivation of proteases by the low pH. This is unlikely, however, as many proteases are naturally located in the lysosome and therefore have a low pH optimum. Another explanation might be that at pH 5.0, the capsid is assembled, and since the N-terminus is localized on the interior of the capsid, it is protected from cleavage by proteases or other degradative processes. It is known from literature[3] that the N-terminal part of the capsid protein is cleaved predominantly at residue 34, resulting in an N-terminally truncated capsid protein (wt CP NΔ34*). *Tang et al.* [3] have shown that this truncated protein is still able to assemble into the native T=3 structure when assembly is induced by lowering the pH to 5.0. Besides T=3 particles, the capsid proteins can also assemble into pseudo T=2 and T=1 particles. These forms are seldomly observed with the full length wt CP.

It has been suggested that the full length wt CP is not able to assemble into T=1 particles, because the formation of a T=1 particle requires a sharper angle between the two capsid proteins of a dimer, which supposedly can only be made by the truncated wt CP. It is clear that by using pH induced assembly conditions formation of the T=1 particle only occurs with the truncated wt CP, but this does not prove that it is essentially impossible to form a T=1 particle from full length wt CP. It could be that formation of a T=1 particle from full length wt CP is possible, but energetically unfavorable at pH 5.0.

Sikkema et al.[4] have shown that particles of the same size as T=1 particles can be formed when PSS is added to the wt CP at pH 7.5. The interaction between the negatively charged PSS and the positively charged capsid protein interior probably provides the driving force for this assembly. The capsid proteins (CP) in these experiments consisted mostly of full length wt CP, so it seems likely that the particles were composed predominantly of this type of CP, which are deemed

* The N in this notation indicates that the amino acids are missing from the N-terminus, Δ indicates missing amino acids, 34 states the number of missing amino acids.

unable to form the sharper angle. Unfortunately, wt CP degrades during the experiments at pH 7.5, and thus it was not possible to form the particles from capsid protein samples that were 100% intact. His-CP seems to degrade at a slower rate than the wt CP (see below), so by using this CP instead of wt CP it might be possible to use a 100% non-degraded capsid protein sample for these assembly experiments.

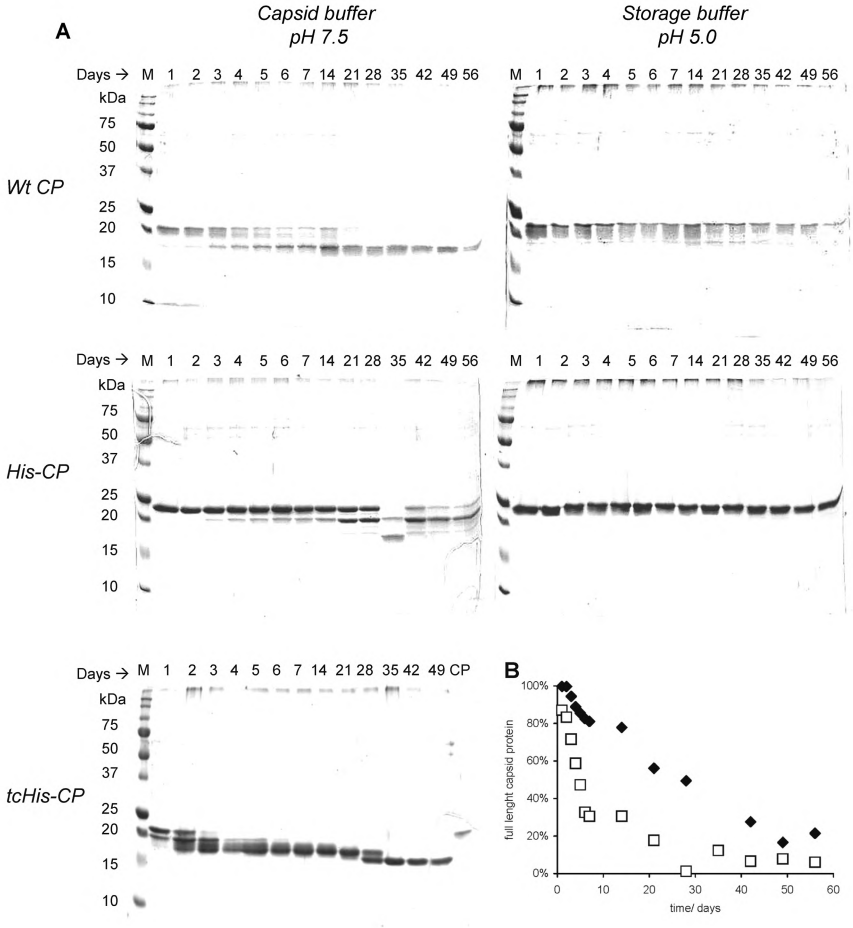

Figure 5 *Degradation of capsid proteins over time. A) SDS-PAGE analysis of degradation over time. tcHis-CP = thrombin cleaved His-CP. B) Percentage of full-length capsid protein at pH 7.5 as function of the time. Filled diamonds represent His-CP, open squares represent wt CP.*

To investigate the difference in degradation rate between wt CP and His-CP, and the effect of pH on the degradation process, both proteins were stored at 4°C at pH 5.0 and pH 7.5 for two months, and were sampled in the course of time. To prevent further degradation of the samples obtained in this way, they were mixed with sample buffer for SDS-PAGE, boiled, and stored at -20°C. This process completely denatures all proteins in the sample, but does not degrade the proteins any further. After two months all collected samples were analyzed by SDS-PAGE (figure 5).

This experiment showed that at pH 5.0 both wt CP and His-CP do not visibly degrade within two months. At pH 7.5, however, both capsid proteins degrade substantially, although the wt CP degrades more rapidly. The degradation rate of wt CP seems to be very fast in the first week and then slows down, whereas His-CP seems to degrade at a more constant rate (figure 5b). It might be that the His-tag provides some degree of protection from degradation. Mass-analysis of the His-CP sample (figure 6) and data from literature[3] revealed that His-CP and wt CP and are not cleaved at the same position.

The main degradation product of wt CP is the truncated protein lacking the first 34 amino acids, while the main degradation product for His-CP is the truncated protein missing the first 21 amino acids, which are exactly the additional amino acids of the His-tag and the thrombin cleavage site. The main degradation product of the His-CP therefore seems to be the wt CP (table 3).

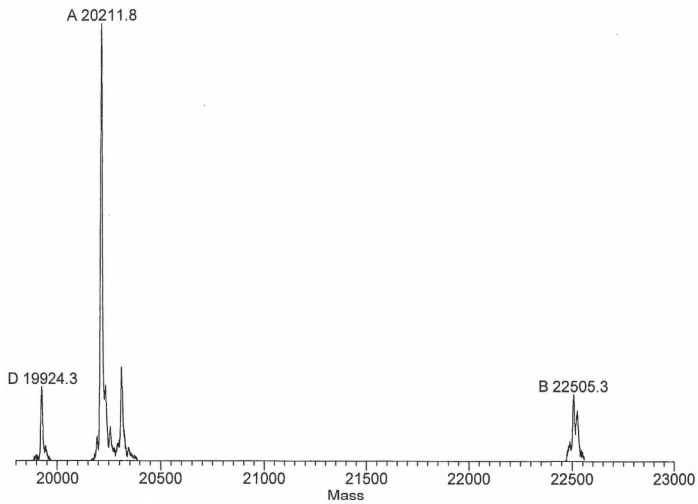

Figure 6 *ESI-TOF mass spectrum of partially degraded His-CP*

Table 3 *Amino acid sequences of the N-termini of His-CP and its main degradation products. The first amino acid of wt CP NΔ34 is highlighted in grey. The first amino-acid of the thrombin cleaved His-CP is printed in bold.*

calculated mass	measured mass	partial amino acid sequence
22506.6	22505.3	GSSHHHHHHSSGLVPR**G**SHMMSTVGTGKLTRAQRRAAARKNKRNTRVVQPVIVEPIAS
20212.1	20211.8	STVGTGKLTRAQRRAAARKNKRNTRVVQPVIVEPIAS
19924.8	19924.3	GTGKLTRAQRRAAARKNKRNTRVVQPVIVEPIAS

This could be very convenient for the study of the assembly of full-length CP into T=1 particles, since even the degraded form of the His-CP is not shorter than the wt CP. Unfortunately, preliminary studies showed that His-CP was not able to form T=1 like particles. Upon incubation with PSS, slightly larger particles were formed, as determined by FPLC analysis. The 21 amino acids longer N-terminus of the His-CP might be the cause of this. The N-terminus of the His-CP can be cleaved at the thrombin cleavage site, resulting in a protein that contains only four amino-acids more than the wt CP. Since it is possible that the His-tag retards the degradation, the degradation rate of the thrombin cleaved His-CP (tcHis-CP) might be different. Therefore, the degradation of the tcHis-CP was studied following the same procedure as outlined for the His-CP and wt CP (figure 5). This protein seems to rapidly degrade first to the wt CP, and then to a truncated version of wt CP, which further degrades to a still more truncated species. The exact size of these truncated versions was not determined. Since it seemed that the His-CP degrades more quickly after addition of thrombin, PSS was added at different time-points after the addition of thrombin. Just before addition of PSS a sample was taken for SDS-PAGE analysis (figure 7a).

In addition to the tcHis-CP samples also full-length His-CP, wt CP and degraded wt CP were tested. The resulting particles were studied on DLS, TEM and FPLC (figure 7b-e). To determine if the protein composition of the formed particles was identical to the initial protein composition, purified FPLC fractions were analyzed on SDS-PAGE, (figure 7a), showing only a slight difference in composition before and after addition of PSS. Full length His-CP was incorporated into the particles to a slightly lesser extent than wt CP in the largest particles (figure 7b). The degraded wt CP almost completely lacked the positively charged N-terminus, which supposedly interacts with the PSS, but still particle formation was observed at pH 7.5.

Aniagyei et al.[5] observed assembly of a genetically truncated NΔ34 CP, which thus lacks the positively charged N-terminus, around an anionic gold nano-particle core. They postulate that the lack of electrostatic interaction between the negatively charged core and the positively charged N-terminus might be compensated by the reduced intersubunit electrostatic repulsion. This might also apply to our PSS containing system. The observed particles formed from the degraded wt CP seemed to be much smaller than all other observed particles, and with an average diameter of about 14 nm, do not seem to fit to any known icosahedral symmetry. The size difference between the particles formed from wt CP and tcHis-CP and His-CP seemed to be much smaller. On TEM almost no difference in size could be detected, while FPLC analysis seemed to indicate a gradual increase in particle size, corresponding to an increased average protein size. DLS also showed the same trend, with the exception of the full length His-CP, which seemed to form smaller particles on DLS. Detailed analysis of the TEM pictures revealed that no 28 nm sized particles were formed. Most particles were in the range of 17-23 nm, which comprises both T=1 and pseudo T=2 particles. Since the difference in size between T=1 and T=2 particles is quite small, it is difficult to distinguish between them. The gradual shift in size might be explained by a shift of the T=1/ T=2 ratio. It is possible that the full-length His-CP is not able to form the curvature needed to form T=1 particles and is therefore kinetically trapped into T=2 particles.

It is striking that on FPLC at least there seems to be a large difference in size between wt CP and non-degraded tcHis-CP (samples 2 and 3). DLS and TEM do not indicate the same size difference, so it is difficult to conclude much from this. In order to really determine the size and structure of the formed particles cryo-TEM measurements and analysis is needed. Therefore, we cannot conclude here that it is possible to form T=1 particles from non-degraded capsid protein, but it seems likely. Especially, since the amount of degradation in the wt CP mostly used for the formation of the PSS particles is usually very low. In order to form a T=1 particle, all dimers need to have a sharper dihedral angle.[2] So if the full-length wt CP would indeed be unable to form this sharper angle, it seems unlikely that a small amount of degradation would suffice for the formation of T=1 particles. The seemingly inability of the larger His-CP to form such a particle does, however, suggest that at least for an elongated N-terminus, it is plausible to assume that the protein dimers are conformationally restricted to assembly in a structure with exclusive T=1 symmetry.

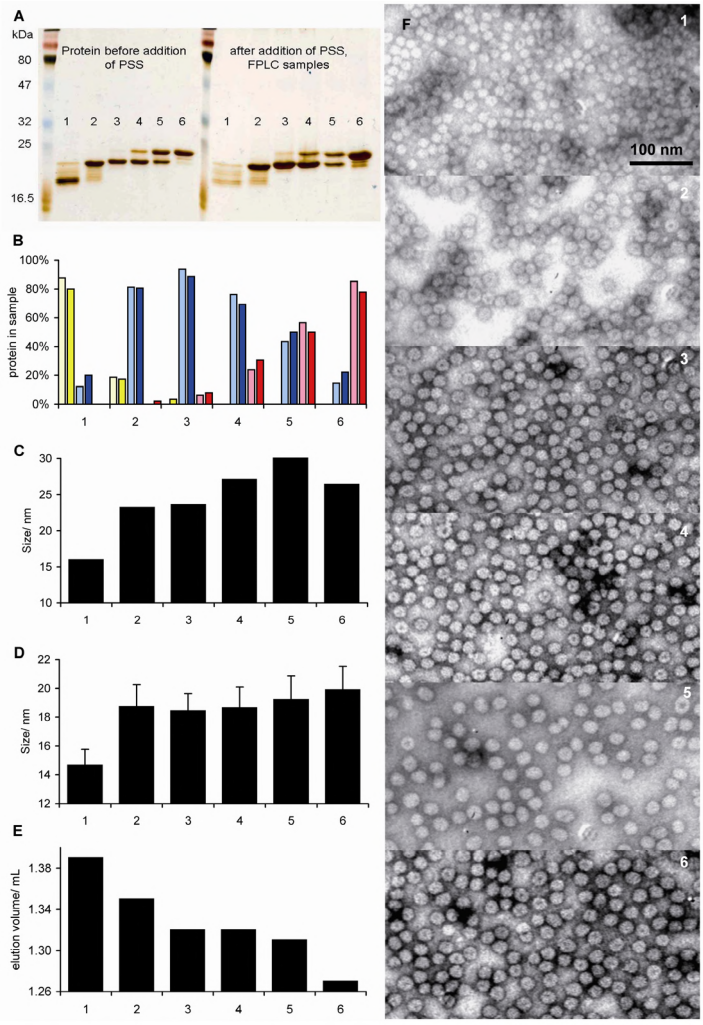

Figure 7 *Addition of PSS to capsid protein of various lengths. A) SDS-PAGE analysis of the samples before addition of PSS and from FPLC purified PSS capsid protein samples. B) Percentage of protein in the samples before addition of PSS (light colors), and in the formed particles after addition of PSS (dark colors) as determined by analysis of the SDS-PAGE gel shown in A. Yellow bars represent degraded wt CP, blue bars represent wt CP, red bars represent His-CP. C) DLS analysis of the formed particles, D) average size of formed particles as determined by TEM, error bars indicate standard deviation. E) Elution volumes of the particles as determined by FPLC analysis, FPLC was equipped with a superose 6 column. F) TEM pictures of the particles (the scale bar applies to all images). 1: degraded wt CP, 2: wt CP, 3: tcHis-CP 210 minutes after thrombin addition, 4: tcHis-CP 90 minutes after thrombin addition, 5: tcHis-CP 30 minutes after thrombin addition, 6: His-CP.*

Experimental section

Methods and Materials

Protein concentrations were determined using a Cary 50 Conc (Varian, Middelburg) UV-VIS spectrophotometer using a quartz cuvette with a path length of 3 mm. FPLC measurements were performed using a Superose 6 PC 3.2/30 analytical column from GE lifesciences, on an Amersham Ettan LC system, fitted with a fractionating device. Buffers for FPLC were filtered with a Millipore 0.2 μM filter before use. TEM grids (Formvar-Carbon) were exposed to an electron discharge treatment using a Cressington Carbon coater and power unit. The sample was applied to the grids by adding a 5 μL drop of sample solution (~ 0.2 mg/mL) to the grid and carefully removing it after 1 min. immersion using a filter paper. The grid was allowed to dry for at least 15 min. before applying 5 μL of a 2% (w/v) uranyl acetate aqueous solution, which was removed after 15 s. The grid was again allowed to dry for at least 15 min. Samples were studied on a JEOL JEM-1010 TEM (Jeol, Japan) instrument. DLS measurements were performed on a Zetasizer Nano S (Malvern Instruments Ltd, England). After purification of the proteins two buffers were used; a buffer of pH 7.5 (50 mM phosphate, 0.5 M NaCl) and a buffer of pH 5.0 (50 mM acetate, 0.5 M NaCl, 10 mM $CaCl_2$), the pH of the buffers was set with HCl.

Creation of a heterologous expression system for the capsid protein in E.coli cells.[2,6]

Detailed procedures for the heterologous expression of CP in *E.coli* will be published in the forthcoming thesis of L.J.H. Hendriks.[2] CCMV virus particles were purified from cowpea plants and RNA was extracted as described in the literature.[7] An amount of 3 μg of RNA was used in a reverse transcriptase reaction using the primer AGTAGGATCC*CTAATACACCGGAGTGAAG* (BamHI restriction site underlined, complementary part italic). The primer is complementary to the cDNA encoding the C-terminus of CP. The obtained cDNA was directly used in a PCR reaction, using the primer described above and a primer complementary to the cDNA encoding the N-terminus of CP: AGTACATATG*ATGTCTACAGTCGGAACAGGG* (*NdeI* restriction site underlined, complementary part italic). A 581 bp fragment was obtained, which was digested with *NdeI* and *BamHI* before it was ligated into a pET-15b *E. coli* expression vector digested with the same restriction enzymes. The obtained plasmid was transformed to *E. coli* and its sequence was confirmed by sequence analysis.

Expression of His-CP

One colony of BL21(DE3)pLysS cells expressing His-CP was used to inoculate 100 mL of LB medium containing ampicillin (0.050 g/L) and chloramphenicol (0.025 g/L). After growth overnight at 30°C this culture was used to inoculate 900 mL of LB medium containing ampicillin (0.05 g/L) and chloramphenicol (0.025 g/L) and grown at 30°C. Protein expression was induced during logarithmic growth (OD_{600} =0.4-0.6) by addition of IPTG to a final concentration of 1 mmol/L.

Purification of His-CP

Cells were harvested after 7 hrs of expression by centrifugation at 4000 rpm for 15 min. at 4°C. The supernatant was discarded and the pelleted cells were stored at -20°C. After thawing, the pelleted cells were resuspended in approximately 10 mL of lysis buffer (50 mM NaH_2PO_4, 10 mM imidazole and 1.5 M NaCl pH 8.0). Approximately 10 mg of lysozyme (EC 3.2.1.17 from Fluka) was added and the solution was incubated at 4°C for 30 min. The solution was sonicated for 5 times 10 seconds with duty cycle 40 and output control 6 (Branson Sonifier 250, marius instruments Nieuwegein, the Netherlands). RNase (10 ug/mL) and DNase (5ug/mL) were added to the solution and the mixture was incubated at 4°C for 15 min. The solution was centrifuged at 10,000 rpm for 20-30 min. to pellet the cellular debris. The supernatant was incubated with 1 mL of Ni-NTA agarose beads for 1 h at 4°C. The flow through was collected and the column was washed with 40 mL of wash buffer (50 mM NaH_2PO_4, 25 mM imidazole and 1.5 M NaCl pH 8.0)). The capsid protein was then eluted from the column using approximately 10 mL of elution buffer (50 mM NaH_2PO_4, 250 mM imidazole and 1.5 M NaCl pH 8.0) and the protein was collected in small fractions. The fractions were dialyzed overnight to buffer pH 7.5, to remove the excess of imidazole. The protein purity was checked with SDS-PAGE gel electrophoresis. Details about the experiments to optimize the protein purification can be found in the text.

Degradation of capsid proteins

Directly after purification of wt CP and His-CP the proteins were divided in two fractions, which were dialyzed to two different buffers: capsid storage buffer pH 5.0 (50 mM NaAc, 1 mM NaN_3, 1 M NaCl), and capsid buffer pH 7.5 (50 mM Tris, 500 mM NaCl, 1 mM EDTA, 10 mM $MgCl_2$). The four samples were diluted to a concentration of 1 mg/mL and stored at 4°C for 2 months. At selected time points, samples were taken from these stock solutions for analysis by SDS-PAGE. Sample loading buffer and a capsid buffer pH 7.5 containing 0.2 M NaCl was added to the samples. The capsid buffer was added to dilute the samples to a final concentration of 0.16 mg/mL, to increase the pH and to decrease the salt concentration of the samples, in order to obtain nicely defined bands on SDS-PAGE. After addition of sample buffer and capsid buffer the samples were boiled for 5 min., and stored at -20°C to prevent further degradation of the proteins. After two months all collected samples were defrosted and analyzed on SDS-PAGE (figure 5).

Formation of PSS stabilized capsids

An amount of 0.44 µL thrombin (Novagen 1.7 U/µL) was added to 1 mg of His-CP (1.85 mg/mL) in capsid buffer pH 7.5. At selected time points samples were isolated from this stock solution. The isolated samples were all treated according to the following procedure; 2 µL of the selected samples were diluted with sample buffer and MQ water to a final concentration of 0.18 mg/mL. The diluted samples were immediately boiled for 5 min. and stored at -20°C. 1 µL PSS dissolved in MQ water (0.32 M) was added to 100 µL sample and

capsids were allowed to form for 15 min. at 4°C. The formed capsids were analyzed on DLS, after which the sample was analyzed by FPLC equipped with a superose 6 column. The capsid fraction eluting from the FPLC at V = ~1.3 mL was analyzed on TEM, and 16 µL of the fractions were mixed with 4 µL sample buffer, boiled for 5 min. and stored at -20°C. All collected samples were analyzed on SDS-PAGE and detected using silverstaining.

References

1. Lavelle, L., et al., *Phase Diagram of Self-assembled Viral Capsid Protein Polymorphs.* J. Phys. Chem. B, **2009**, *113*, 3813.
2. Forthcoming thesis of Hendriks, L.J.A., Radboud University Nijmegen.
3. Tang, J.H., et al., *The role of subunit hinges and molecular "switches" in the control of viral capsid polymorphism.* J. Struct. Biol., **2006**, *154*, 59.
4. Sikkema, F.D., et al., *Monodisperse polymer-virus hybrid nanoparticles.* Org. Biomol. Chem., **2007**, *5*, 54.
5. Aniagyei, S.E., et al., *Synergistic Effects of Mutations and Nanoparticle Templating in the Self-Assembly of Cowpea Chlorotic Mottle Virus Capsids.* Nano Lett., **2009**, *9*, 393.
6. Minten, I.J., et al., *Controlled Encapsulation of Multiple Proteins in Virus Capsids.* J. Am. Chem. Soc., **2009**, *131*, 17771.
7. Bujarski, J.J., *Bromovirus Isolation and RNA Extraction.* Vol. 81. 1998. 183.

Chapter 6

*Controlled encapsulation of multiple proteins in virus capsids**

Introduction

Viruses are masters in encapsulating nucleic acids; they are typically composed of a shell of highly organized protein molecules, which surround densely packed DNA or RNA chains. Recently, these protein shells (also called capsids) have also been used to package other guest compounds.[1-4] CCMV is particularly suited for this purpose, due to its special assembly properties, which makes it possible to encapsulate compounds by mixing them with CCMV dimers at pH 7.5 and subsequently lowering the pH to entrap material inside the capsid.[1,5]

Since the CCMV capsid interior is positively charged, encapsulation works especially well with negatively charged guest molecules like inorganic salts and negatively charged polymers as described in chapter 3.[1,6] However, potentially useful cargos like enzymes are usually not negatively charged at acidic pH, resulting in a low encapsulation efficiency of these enzymes. Low encapsulation efficiency is not always a disadvantage, as has been shown in our study of a virus-based enzyme nanoreactor[5] which has provided new insight in the behavior of single enzyme molecules.

Following up on this work, it is of great interest, to study the behavior of multiple enzymes in the confined space of a virus nanoreactor. For example, cascade reactions catalyzed by a series of enzymes might take place more efficiently. In addition, such systems may provide new insights into the way living cells work, since viral nanoreactors could serve as a model system for simple cell organelles.

* This work was published: *JACS*, **2009**, 131, 17771

In order to achieve this goal, we need to improve and control the encapsulation efficiency of enzymes in the CCMV capsid, compared to the statistical encapsulation strategy previously reported by us.[5]

In this chapter we describe a novel encapsulation method, which makes use of a non-covalent anchoring moiety to attach the target protein to the capsid protein prior to assembly of the latter. In this way the guest encapsulation efficiency is significantly improved, while to a large extent control over the amount of encased proteins is obtained.

Results and Discussion

Non-covalent anchoring was chosen over covalent anchoring methods, because it eliminates the need for chemical reactions, which are often not bio-orthogonal, to bind the two proteins together. As a non-covalent anchor, we used a heterodimeric coiled-coil protein[7]. Coiled-coil motifs are oligomerization domains frequently found in nature[8,9] and are also used by man, for example in affinity chromatography,[10] biosensors,[11] hydrogels[12], and in the study of liposome fusion.[13] Coiled-coil motifs consist of a seven-amino acid-residue repeat, denoted as $\boldsymbol{(abcdefg)_n}$, in which residues *a* and *d* are usually hydrophobic amino acids, which align when the peptide sequence folds into a helical conformation. A hydrophobic strip is subsequently formed at one side of the coiled-coil, which is the primary driving force for the self-assembly of two of such species.[14] While nature commonly uses homodimeric coiled-coils, we chose a heterodimeric coiled-coil to avoid the problem of self-dimerization of either the capsid protein or the guest protein. To minimize the risk that the the introduction of a coiled-coil would alter the structure and assembly properties of the capsid protein, a short coiled-coil was chosen, which nonetheless still has a small dissociation constant (K_d= 7×10^{-8} M) (figure 1).[7]

Heterodimerization can be achieved by the placement of charged residues (glutamic acid or lysine) at the *e* and *g* positions of the coiled-coil motif. The resulting coiled-coil with glutamic acid at these positions is called the E-coil and the one with lysine at these positions, is called the K-coil. The complete amino acid sequences of these peptides are $(EIAALEK)_3$ for the E-coil, and $(KIAALKE)_3$ for K-coil.

Since the capsid interior is positively charged, we chose to attach the positively charged K-coil to the capsid protein ensuring that the coiled-coil does not

stick to the capsid protein surface, which presumably would make it less available for binding with the E-coil. In order to attach the K-coil to the capsid protein, we cloned the gene encoding for the coiled-coil sequence in an expression plasmid containing the capsid protein gene, introducing the K-coil at the N-terminus of this protein. From this point on this protein will be referred to as CK.

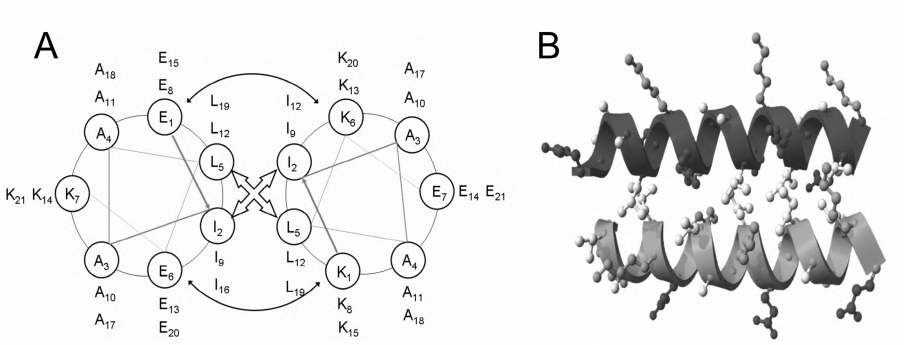

Figure 1 *Schematic representations of the E/K coil heterodimer. A) Helical wheel representation of the E/K coil heterodimer. The helix propagates into the page from the N to the C-terminus. Hydrophobic interactions are depicted with open arrows, ionic interactions with closed two sided arrows. The left wheel represents the E-coil with amino acid sequence* $(EIAALEK)_3$ *and the right wheel represents the K-coil with amino acid sequence* $(KIAALKE)_3$*. B) Cartoon of the E/K coil heterodimer. The E-coil is represented in dark grey, and the K-coil is represented in grey. Negatively charged amino-acids are depicted in dark grey, positively charged amino acids are depicted in grey and hydrophobic amino acids are depicted in light grey.*

The N-terminus was chosen for modification since the C-terminal part of the capsid protein is involved in the dimerization and assembly of the capsid proteins. Since the E-coil and K-coil bind in a parallel fashion, the E-coil was introduced at the C-terminus of EGFP by introducing it in an expression plasmid containing the gene sequence for EGFP and an N-terminal His-tag sequence. From this point on this protein will be referred to as GE. The modified proteins were brought to expression in *E.coli* cells. In order to assemble and purify the EGFP-capsid protein complex (GE-CK complex), we used the N-terminal His-tag sequence of GE to immobilize it on a nickel-NTA column. The E-coil is attached to the C-terminus of GE, and therefore available for binding to the modified capsid protein, which was added in excess to the column containing bound GE (figure 2).

The modified capsid protein was allowed to bind for at least an hour before washing away all non-bound protein and eluting the complex from the column with an excess of imidazole. To remove the imidazole the complex was immediately dialyzed to a pH 7.5 buffer. In principle only the complete GE-CK complex will elute from the column, however, small impurities and any GE protein that did not bind to a capsid protein could have co-eluted as well from the column.

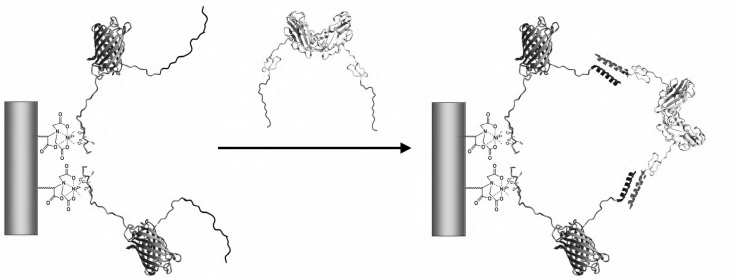

Figure 2 *Purification of the GE-CK complex. Bacterial lysate containing GE is added to a nickel-NTA column. Only GE will bind to the nickel-NTA with its N-terminal His-tag. A wash step removes all other proteins that lack the His-tag. The lysate containing the capsid protein is then added. CK binds to the C-terminal coiled-coil of GE with its N-terminal coiled-coil. After another wash step, the entire complex is eluted from the column, using an excess of imidazole.*

The eluate was therefore further purified on a FPLC system equipped with a Superdex 200 column, and the identity and purity of the proteins was assessed by SDS-PAGE analysis and immunoblotting (figure 3).

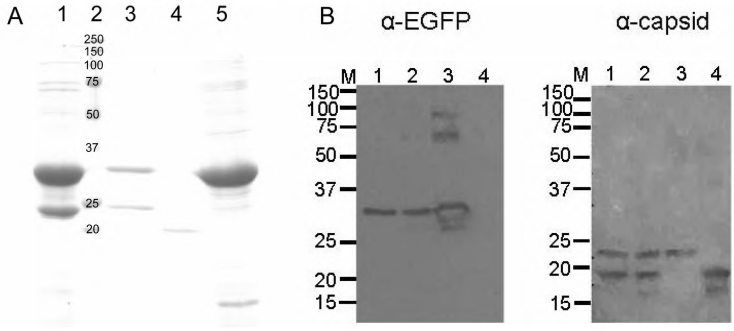

Figure 3 *Analysis of identity and purity of proteins used for EGFP encapsulation A) SDS-PAGE gel, lane 1: elution fraction of GE-CK complex from Ni-NTA column. Lane 2: protein marker, numbers represent molecular weights in kDa. Lane 3: GE-CK complex after purification by FPLC. Lane 4: wt CP. Lane 5: elution fraction of GE from Ni-NTA column.*

B) Immuno blots. Lane 1 and 2: 1.1 mL peak fractions from the FPLC of capsids with EGFP encapsulated at pH 5.0. Lane 3: GE-CK complex at pH 7.5. Lane 4: wt CP. Left blot was probed with anti-EGFP antibody, right blot was probed with anti-capsid antibody.

The ratio with which GE is bound to CK could be determined by comparing the UV-Vis spectra of pure GE with the spectra of the GE-CK complex (figure 4).

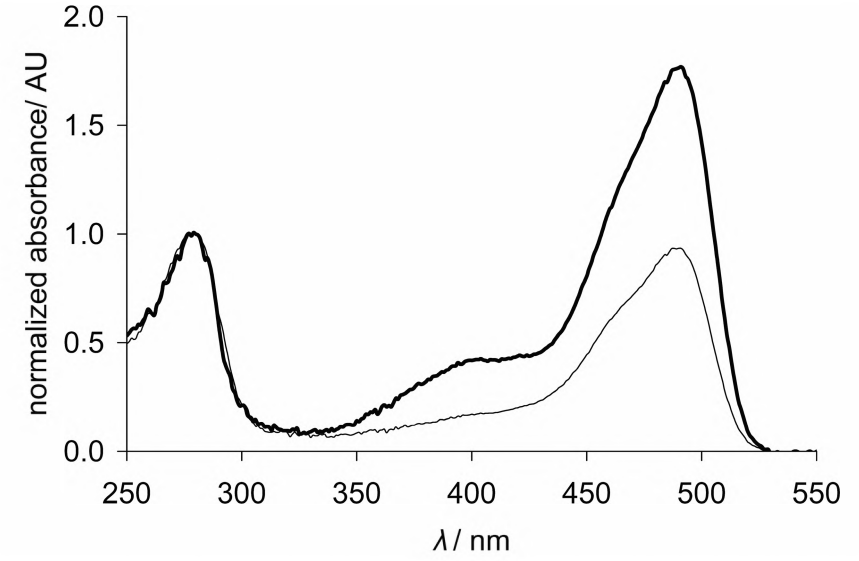

Figure 4 *UV-Vis spectra of GE (thick line) and GE-CK complex (thin line) at pH 7.5. Both absorption spectra are normalized to an absorbance of 1.0 at λ=280 nm.*

Besides the normal protein absorption at λ=280 nm, EGFP also has a specific absorption at λ=490 nm at pH 7.5, which is approximately twice as intense as the λ=280 nm absorption. However, when GE was bound to CK, which does not absorb at λ=490 nm, the λ=490 nm absorption intensity was about equal to the λ=280 nm absorption. This suggests that GE and CK are approximately bound in a 1:1 ratio. (For a detailed calculation of this ratio see experimental section). Because the capsid protein always exists as a dimer, it is probable that a complex of a capsid protein dimer bound to two EGFP proteins is formed. This was further confirmed by FPLC analysis of the complex. At low concentrations the GE-CK complex eluted from the FPLC column at V=1.4 mL, which is the expected elution volume for a complex of this molecular weight (figure 5).

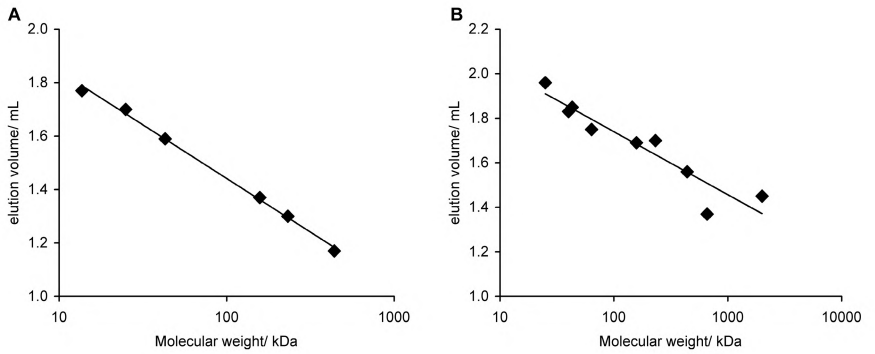

Figure 5 *Calibration curves for the FPLC system. Diamonds represent measured proteins. Line represents fitted calibration curve. A) Calibration curve of FPLC equipped with a superose 200 column. B) Calibration curve of FPLC equipped with a superose 6 column.*

In order to encapsulate GE, the GE-CK complex was mixed with wild-type capsid protein, isolated from the CCMV virus grown in plants. This was done in various ratios, to allow control over the amount of encapsulated GE and to prevent overcrowding of capsid. After a short incubation period this mixture was dialyzed to pH 5.0 (figure 6) and analyzed by FPLC (figure 7). The amount of GE encapsulated per capsid can be calculated using the 280 nm/395 nm absorbance ratio of the capsid peak at V=1.1 mL.

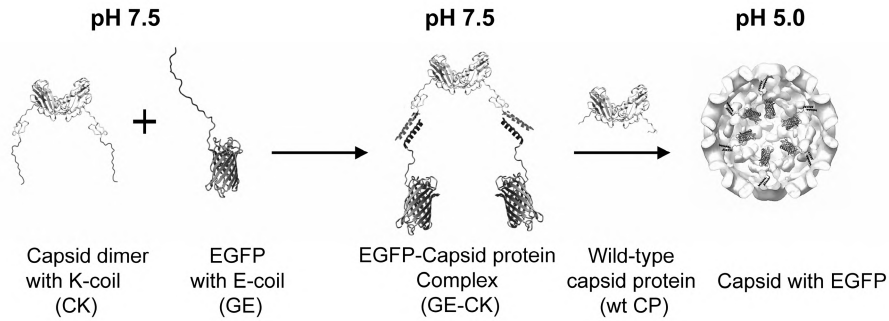

Figure 6 *Schematic representation of EGFP encapsulation. The GE-CK complex is mixed with wild-type capsid at pH 7.5 and subsequently dialyzed to pH 5.0 to induce capsid formation.*

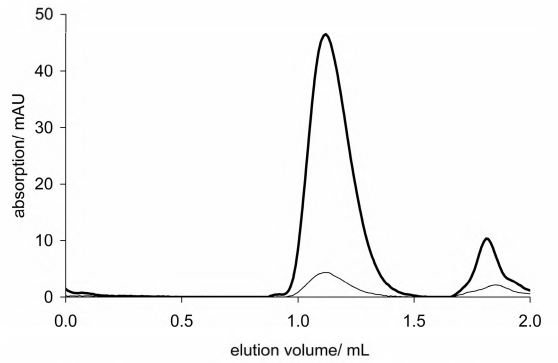

Figure 7 *FPLC trace of a mixture of GE-CK complex with wild-type capsid at pH 5.0. The FPLC system is equipped with a Superose 6 column. The thick line represents the protein absorption at λ=280 nm, and the thin line the EGFP specific absorption at λ=395 nm. The capsid with encapsulated GE elutes at V=1.1 mL, and the unassembled capsid dimers and GE-CK complex elute at ~1.8 mL.*

Since GE absorbs at both λ=280 and λ=395 nm, at pH 5.0, and CK only at λ=280 nm, the CK/GE ratio was determined from these absorption ratios using the respective extinction coefficients *(for determination of the extinction coefficients and calculation of the number of encapsulated GE molecules see experimental section).*

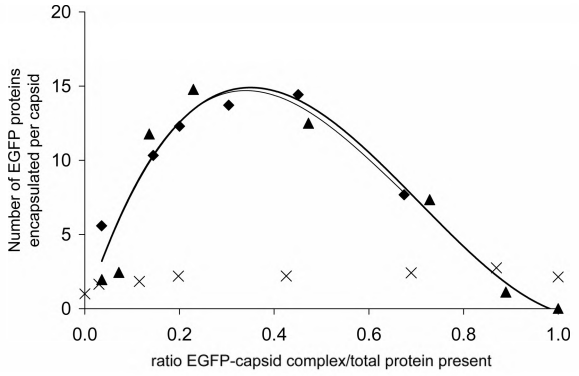

Figure 8 *Number of encapsulated EGFP proteins per capsid as a function of the GE-CK complex to total protein ratio. Diamonds and triangles represent data points of duplicate experiments. Crosses represent negative control experiments in which unbound GE was used instead of the GE-CK complex. The thick line represents the polynomial trend line through the data points depicted with a triangular shape; the thin line represents the polynomial trend line through the diamond-shaped data points.*

The data show that it is possible to encapsulate up to 15 EGFP proteins per capsid (figure 8), and that the amount of encapsulated EGFP to a certain extent can be controlled by varying the ratio between the wt CP and the GE-CK complex. The experiments proved to be highly reproducible.

Throughout this chapter the assumption is made that EGFP is located on the capsid interior, this is based on the following reasoning; In the crystal structure of the CCMV virus, the N-termini of the capsid proteins are situated on the inside of the capsid. Since EGFP is attached to the N-terminus of the capsid protein, it would seem likely that upon assembly of the capsid, EGFP will be situated on the inside of the capsid. This a priori assumption is further supported by TEM micrographs (figure 9).

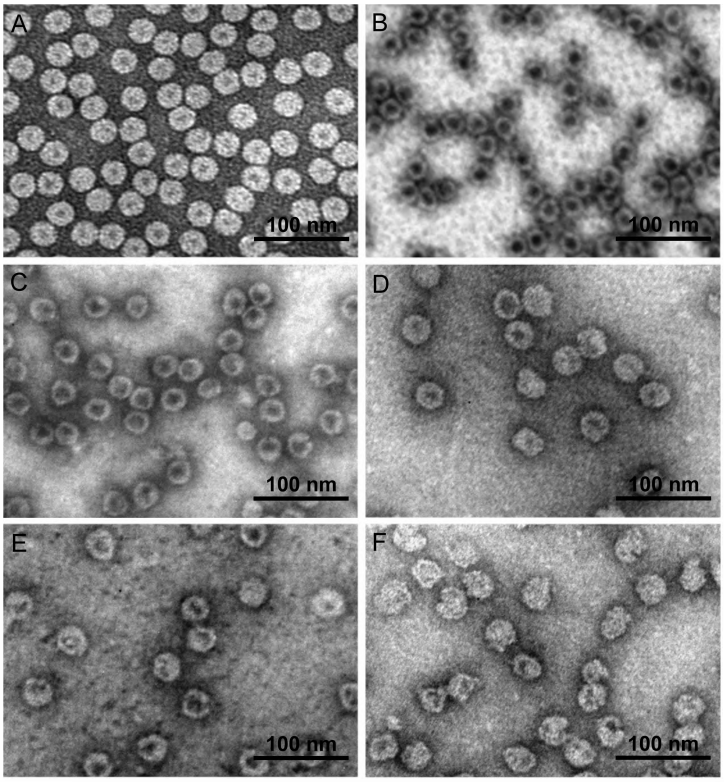

Figure 9 *TEM micrographs of uranyl acetate stained samples. A) CCMV virus. B) Wild-type capsids without EGFP. C) Capsids filled with 2 EGFP per capsid D) Capsids filled with 7 EGFP per capsid. E) Capsids filled with 12 EGFP per capsid. F) Capsids filled with 15 EGFP per capsid.*

In order to visualize capsids by TEM, a negative staining agent, uranyl acetate was used. The protein is not stained and thus will appear lighter than the background. In this way, the CCMV virus looks like a white filled circle, whereas the empty capsid has a black spot in the center of the circle. This is caused by the accumulation of uranyl acetate inside the empty cavity of the capsid. TEM pictures of capsid filled with different amounts of EGFP show that the staining of the capsid interior decreases as the amount of encapsulated EGFP increases. This indicates that the capsid is no longer empty, and thus that the EGFP is encapsulated on the inside of the capsid.

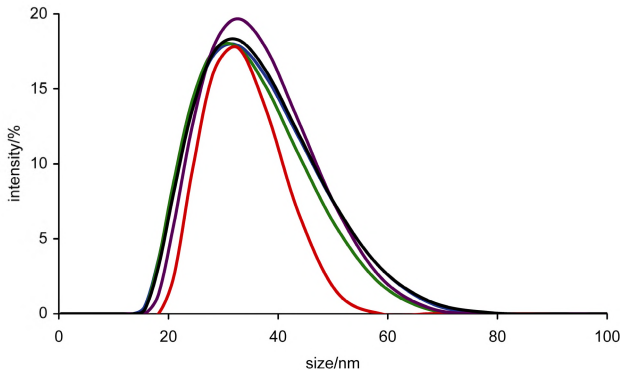

Figure 10 *DLS plot of capsids with and without EGFP. Black line: CCMV, Blue line: capsids filled with 2 EGFP per capsid (negative control, statistical encapsulation of GE). Purple line: capsids filled with 12 EGFP per capsid. Green line: capsids filled with 13 EGFP per capsid. Red line: capsids filled with 14 GE per capsid.*

Furthermore, DLS data showed that the particle size did not increase when EGFP proteins were encapsulated (figure 10). If the proteins would be attached to the outside of the capsid instead of the inside, an increase in capsid size would be expected. Finally, anti-EGFP antibodies did not appear to bind to the capsids with encapsulated GE. This indicates that no EGFP is present on the outside of the capsids. Experimentally this was tested by incubating both empty capsids and capsids loaded with 7 EGFP per capsid with anti-EGFP antibody labeled with rhodamine. These mixtures were analyzed by FPLC equipped with a Superose 6 column and compared with samples to which no anti-EGFP was added. This experiment showed that the anti-EGFP specific absorption from the rhodamine at λ=516 was the same for all samples, indicating that no antibody was bound to the EGFP filled capsids (figure 11).

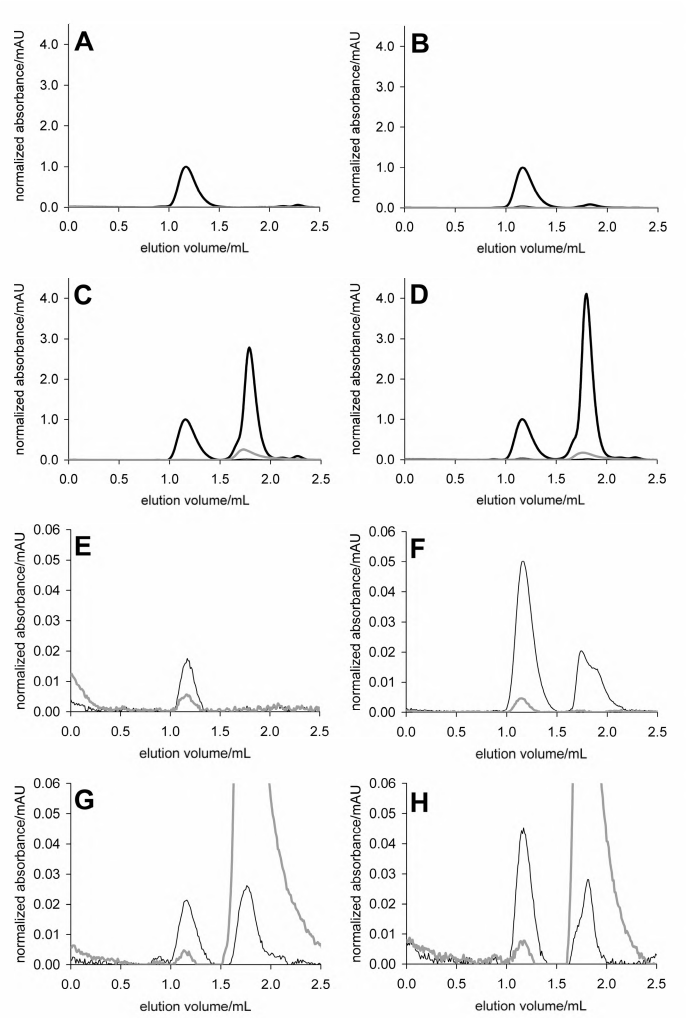

Figure 11 *FPLC graphs of empty capsid and capsid with encapsulated GE, with and without anti-EGFP antibodies. Thick black lines represent protein absorption at λ=280 nm, thin black lines represent EGFP specific absorption at λ=395 nm and thick grey lines represent the absorption of the rhodamine labeled anti-EGFP antibody at λ=516 nm. All graphs have been normalized to an absorbance of 1.0 at the capsid peak (1.1 mL) at λ=280 nm. Graphs E-H are enlargements of graphs A-D, for clarity the λ=280 nm absorption is not depicted in graphs E-H. The capsids elute at 1.1 mL, the unbound anti-EGFP antibody elutes at 1.8 mL. Graph A and E show FPLC traces of the empty capsid. Graphs B and F show FPLC traces of capsid with encapsulated EGFP. Graphs C and G show FPLC traces of empty capsids incubated with anti-EGFP antibody. Graphs D and H show FPLC traces of capsid with encapsulated EGFP incubated with anti-EGFP antibody.*

To check if the antibodies were able to bind to GE in buffer pH 5.0, it was also tried to incubate GE with the antibodies and analyze the sample by FPLC. Unfortunately, it appeared very difficult to get a good separation between the unbound GE proteins and the GE-antibody complex with the help of FPLC, even when using a Superdex 200 column instead of the Superose 6 column. This was largely due to the fact that the GE proteins seemed to dimerize at very low concentrations at pH 5.0, thus making the difference in size between unbound GE and anti-EGFP antibody too small to result in well separated peaks on FPLC. Therefore, the binding of the anti-EGFP antibodies to the GE proteins was tested by immunoblotting. In order to mimic the conditions used for the GE filled capsids, the immunoblotting experiment was performed at pH 5.0 (for exact procedure see the experimental section). This revealed that the antibody was indeed able to bind to GE at pH 5.0 (figure 12). The above experiments all indicate that GE is situated on the inside of the capsids and not on the outside.

Figure 12 *Immunoblot experiment of GE and wild-type capsid protein probed with anti-EGFP antibody in buffer pH 5.0. Circles indicate the positions of the spots.*

Conclusion

From the data presented in this chapter we may conclude that it is possible to bind proteins inside the CCMV capsid in an efficient and controlled fashion, opening the way to the encapsulation of different types of enzymes in one capsid. Our study furthermore shows that viral capsid assembly is a powerful tool for the controlled formation of bio- nano-structures with diverse functions.

Experimental section

Methods and materials

Protein concentrations were determined using a Cary 50 Conc (Varian, Middelburg) UV-VIS spectrophotometer using a quartz cuvet with a path length of 3 mm.

FPLC measurements were performed using either a Superose 6 PC 3.2/30 analytical column or a Superdex 200 PC 3.2/30 analytical column from GE lifesciences, on an Amersham Ettan LC system, fitted with a fractionating device. Buffers for FPLC were filtered with a Millipore 0.2 μM filter before use. TEM grids (Formvar-Carbon) were exposed to an electron discharge treatment using a Cressington Carbon coater and power unit. The sample was applied to the grids by adding a 5 μL drop of sample solution (~ 0.2 mg/mL) to the grid and carefully removing it after 1 min. immersion using a filter paper. The grid was allowed to dry for at least 15 min. before applying 5 μL of a 2% (w/v) uranyl acetate aqueous solution, which was removed after 15 s. The grid was again allowed to dry for at least 15 min.. Samples were studied on a JEOL JEM-1010 TEM (Jeol, Japan).

DLS measurements were performed on a Zetasizer Nano S (Malvern Instruments Ltd, England), Samples were first purified by FPLC and subsequently measured.

Restriction enzymes and Antartic Phosphatase were obtained from New England Biolabs. DNA containing the coiled-coil sequences was obtained from Biolegio (Nijmegen, The Netherlands).

Two buffers were used, a buffer pH 7.5 (0.5 M NaCl, 0.05 M Tris-HCl, 0.01 M $MgCl_2$ and 0.001 M EDTA) and buffer pH 5.0 (0.5 M NaCl, 0.05 M $NaCH_3COO$, 0.01 M $MgCl_2$ and 0.001 M EDTA), the pH of the buffers was set with HCl.

Creation of a heterologous expression system for the capsid protein in E. coli cells

CCMV virus particles were purified from cowpea plants and RNA was extracted[15]. An amount of 3 μg of RNA was used in a reverse transcriptase reaction using the primer AGTAGGATCC*CTAATACACCGGAGTGAAG* (BamHI restriction site underlined, complementary part italic). The primer is complementary to the cDNA encoding the C-terminus of CP. The obtained cDNA was directly used in a PCR reaction, using the primer described above and a primer complementary to the cDNA encoding the N-terminus of CP: AGTACATATG*ATGTCTACAGTCGGAACAGGG* (*NdeI* restriction site underlined, complementary part italic). A 581 bp fragment was obtained, which was digested with *NdeI* and *BamHI* before it was ligated into a pET-15b *E. coli* expression vector digested with the same restriction enzymes. The obtained plasmid was transformed to *E. coli* and its sequence was confirmed by sequence analysis.

Capsid protein with K-coil construction in vector pET-15b

The insert containing the K-coil DNA sequence was designed to have a 5' NcoI and a 3' NdeI restriction site overhang after annealing of the strands (table 1). The pET-15b vector

containing the capsid protein sequence was digested sequentially with NcoI and NdeI, dephosphorylated with Antartic Phosphatase, and purified by agarose gel elecrophoresis. The vector was mixed with the annealed insert and ligated with T4 DNA ligase. The resulting plasmids were transformed into *E. coli* XL1-Blue cells, the DNA was extracted and the sequence of the capsid protein with K-coil was confirmed by DNA sequencing. The plasmids were then transformed into *E. coli* BL21(DE3)pLysS cells, which were used for production of CK.

GE construction in vector pET-15b

The insert containing the E-coil DNA sequence was designed to have a 5' BsrGI and a 3' BamHI restriction site overhang after annealing of the strands (table 1).

The pET-15b vector containing the EGFP protein sequence was digested sequentially with BsrGI and BamHI, dephosphorylated with Antartic Phosphatase, and purified by agarose gel elecrophoresis. The vector was mixed with the annealed insert and ligated with T4 DNA ligase.

The resulting plasmids were transformed into *E. coli* XL1-Blue cells, the DNA was extracted and the sequence of the EGFP with E-coil was confirmed by DNA sequencing. The plasmids were then transformed into *E. coli* BLR (DE3 pLysS) cells, which were used for production of GE.

Expression of CK

One colony of BL21(DE3)pLysS cells expressing CK was used to inoculate 100 mL of LB medium containing ampicilin (0.050 g/L) and chloramphenicol (0.025 g/L). After growth overnight at 30°C this culture was used to inoculate 900 mL of LB medium containing ampicilin (0.05 g/L) and chloramphenicol (0.025 g/L) and grown at 30°C. Protein expression was induced during logarithmic growth (OD_{600} =0.4-0.6) by addition of IPTG to a final concentration of 1 mmol/L.

Expression of GE

One colony of BLR (DE3) pLysS cells expressing GE was used to inoculate 60 mL of LB medium containing ampicilin (0.05 g/L) and tetracycline (0.0125 g/L). After growth overnight at 30°C this culture was used to inoculate 540 mL of LB medium containing ampicilin (0.05 g/L) and tetracycline (0.0125 g/L) and grown at 30°C. Protein expression was induced during logarithmic growth (OD_{600} =0.4-0.6) by addition of IPTG to a final concentration of 1 mmol/L.

Table 1 *DNA and amino acid sequences. DNA sequences for the inserts used for the creation of the CK and GE proteins. Overhangs are complementary with the overhangs in the vector sequences, created with restriction enzymes. Amino acid sequences for wt CP, CK and GE protein. K-coil and E-coil amino acid sequences are printed in bold lettering.*

First line: K-coil forward (5'→3'), second line: K-coil reverse (3'→5')

```
CATGAAGATCGCCGCCCTGAAGGAGAAGATCGCCGCCCTGAAGGAGAAGATCGCCGCCCTGAAGGAGGG
    TTCTAGCGGCGGGACTTCCTCTTCTAGCGGCGGGACTTCCTCTTCTAGCGGCGGGACTTCCTCCCAT
```

First line: E-coil forward (5'→3'), second line: E-coil reverse (3'→5')

```
GTACACGCTAGCTGAAATTGCTGCTCTTGAAAAAGAAATTGCTGCTCTTGAAAAAGAAATTGCTGCTCTTGAAAAATAAGC
    TGCGATCGACTTTAACGACGAGAACTTTTTCTTTAACGACGAGAACTTTTTCTTTAACGACGAGAACTTTTTATTCGCTAG
```

Wt CP amino acid sequence

```
MSTVGTGKLTRAQRRAAARKNKRNTRVVQPVIVEPIASGQGKAIKAWTGYSVSKWTASCAAAEAKVTSAITISLPNELSSERNKQ
LKVGRVLLWLGLLPSVSGTVKSCVTETQTTAAASFQVALAVADNSKDVVAAMYPEAFKGITLEQLAADLTIYLYSSAALTEGDVI
VHLEVEHVRPTFDDSFTPVY
```

CK amino acid sequence

```
MKIAALKEKIAALKEKIAALKEGMMSTVGTGKLTRAQRRAAARKNKRNTRVVQPVIVEPIASGQGKAIKAWTGYSVSKWTASCAA
AEAKVTSAITISLPNELSSERNKQLKVGRVLLWLGLLPSVSGTVKSCVTETQTTAAASFQVALAVADNSKDVVAAMYPEAFKGIT
LEQLTADLTIYLYSSAALTEGDVIVHLEVEHVRPTFDDSFTPVY
```

GE amino acid sequence

```
MGSSHHHHHHSSGLVPRGSHMLEKREAEAGRLGAGGPVATMVSKGEELFTGVVPILVELDGDVNGHKFSVSGEGEGDATYGKLTL
KFICTTGKLPVPWPTLVTTLTYGVQCFSRYPDHMKQHDFFKSAMPEGYVQERTIFFKDDGNYKTRAEVKFEGDTLVNRIELKGID
FKEDGNILGHKLEYNYNSHNVYIMADKQKNGIKVNFKIRHNIEDGSVQLADHYQQNTPIGDGPVLLPDNHYLSTQSALSKDPNEK
RDHMVLLEFVTAAGITLGMDELYTLAEIAALEKEIAALEKEIAALEK
```

Purification of GE and CK

E. coli cells containing the GE proteins, and *E. coli* cells containing the CK protein, were separately harvested after 5 h of expression by centrifugation at 4000 g for 15 min. at 4°C. The supernatant was discarded and the pelletted cells were stored at -20°C. After thawing, the pelletted cells were resuspended in approximately 10 mL lysis buffer (50 mM NaH_2PO_4, 10 mM imidazole and 300 mM NaCl pH 8.0 for GE, and 50 mM NaH_2PO_4, 10 mM imidazole and 1500 mM NaCl pH 8.0 for CK). Approximately 10 mg of lysozyme (EC 3.2.1.17 from Fluka) was added and the solutions were incubated at 4°C for 30 min.. The solutions were sonicated for 5 times 10 s with duty cycle 40 and output control 6 (Branson Sonifier 250, marius instruments Nieuwegein, the Netherlands). RNase (10 µg/mL) and DNase (5 µg/mL) were added only to the solution containing CK and the mixture was incubated at 4°C for 15 min. The solutions were centrifuged at 10,000 rpm for 20-30 min. to pellet the cellular debris. The supernatant of the GE containing cells were incubated with 1 mL of Ni-NTA agarose beads for 1 h at 4°C. The flow-through was collected and the column

was washed with 20 mL of wash buffer (50 mM NaH_2PO_4, 12.5 mM imidazole and 300 mM NaCl pH 8.0). The supernatant of the CK containing cells was then added to the Ni-NTA beads containing the bound GE protein and incubated for 1 h at 4°C, to allow binding of the E- and K-coil of the GE and CK proteins respectively. The flow-through was again collected and the column was washed with 20 mL of wash buffer (50 mM NaH_2PO_4, 12.5 mM imidazole and 1500 mM NaCl pH 8.0). The GE-CK complex was then eluted from the column using approximately 10 mL of elution buffer (50 mM NaH_2PO_4, 250 mM imidazole and 1500 mM NaCl pH 8.0) and the protein complex was collected. The protein complex was dialyzed overnight to buffer pH 7.5, to remove the excess of imidazole. The protein purity was checked with SDS-PAGE analysis (figure 3)

Production and purification of wt CP

The purification of the CCMV virus and the removal of its RNA were carried out according to literature procedures[16,17].

FPLC purification of GE-CK complex

The GE-CK complex was further purified by FPLC equipped with a Superdex 200 column to remove any non-complexed GE, CK or other proteins. Pure GE eluted at V=1.6 mL as detected at λ=280 and 490 nm, whereas the CK-GE complex eluted at V=1.2-1.4 mL depending on the concentration of the complex.

EGFP encapsulation experiments

CK-GE complex in buffer pH 7.5 (50 mM Tris-HCl, 500 mM NaCl, 10 mM $MgCl_2$, 1 mM EDTA pH 7.5) was added in different ratios to wild type capsid protein in the same buffer (table 2). The proteins were allowed to mix for 5 min., before dialyzing the mixture overnight to buffer pH 5.0. The mixtures were then analyzed on an FPLC system equipped with a Superose 6 column (figure 6). Peak fractions from the FPLC of the capsid with EGFP encapsulated were analyzed by immuno blotting (figure 3), showing intact EGFP-with E-coil, as well as wt CP and capsid protein with K-coil. Anti-capsid antibody was kindly donated by BJM Verduin, Wageningen University.

Calibration curves of FPLC columns

Calibration curves were made for the superdex 200 and superose 6 column using a protein based low molecular weight and high molecular weight gel filtration calibration kit from Amersham biosciences.

Table 2 *Data of the encapsulation experiments*

Exp. #	M wt CP in samples	M GE-CK complex in samples	[a]Ratio [GE-CK complex]/ [total protein]	Number of EGFP proteins encapsulated per capsid
1	7.71×10^{-5}	2.50×10^{-6}	0.03	2.0
	7.39×10^{-5}	4.99×10^{-6}	0.06	2.4
	6.83×10^{-5}	9.36×10^{-6}	0.12	11.8
	6.02×10^{-5}	1.56×10^{-5}	0.21	14.8
	4.02×10^{-5}	3.12×10^{-5}	0.44	12.5
	2.01×10^{-5}	4.68×10^{-5}	0.70	7.3
	8.03×10^{-6}	5.62×10^{-5}	0.87	1.1
	0	6.24×10^{-5}	1.00	0.0
2	7.84×10^{-5}	2.52×10^{-6}	0.03	5.6
	6.86×10^{-5}	1.01×10^{-5}	0.13	10.3
	6.37×10^{-5}	1.39×10^{-5}	0.18	12.3
	5.47×10^{-5}	2.08×10^{-5}	0.28	13.7
	4.25×10^{-5}	3.03×10^{-5}	0.42	14.4
	2.45×10^{-5}	4.41×10^{-5}	0.64	7.7

[a] *Ratio was calculated by dividing molarity of the GE-CK complex by the molarity of the GE-CK complex+M wt CP*

Determination of EGFP extinction coefficients

EGFP was expressed and purified using the same procedure as for the GE protein, but without the addition of CK protein to the bound EGFP. The obtained EGFP was subsequently purified using an FPLC system equipped with a Superdex 200 column. The pure EGFP was concentrated in a 0.5 M NaCl solution and diluted to six different dilutions in a 0.5 M NaCl solution. These stock solutions were then diluted ten times with either buffer pH 7.5 or buffer pH 5.0. From each different combination three samples were made. Each sample was then measured twice with the help of a UV-Vis spectrophotometer. The absorption values (λ=280 and 490 nm in buffer pH 7.5 and 280 and 395 nm in buffer pH 5.0) of the duplicate measurement were averaged. A calibration curve could be determined from the resulting graph (figure 13). Using the calibration curve several extinction coefficients of EGFP were calculated, using the extinction coefficient value of 55000[18] at λ=490 nm at pH 7.5 as a reference value (table 3).

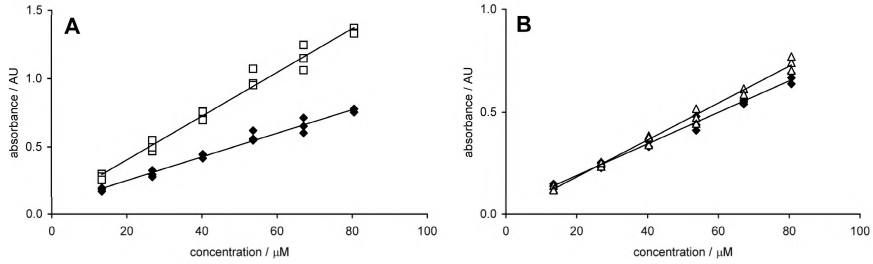

Figure 13 *Calibration curves for the determination of the extinction coefficients of EGFP in buffer pH 7.5 and buffer pH 5.0. A) Calibration curve of EGFP in buffer pH 7.5, open squares represent absorption of the EGFP samples measured at λ=490 nm, filled diamonds represent absorption of the EGFP samples measured at λ=280 nm. B) Calibration curve of EGFP in buffer pH 5.0, open triangles represent absorption of the EGFP samples measured at λ=395 nm, filled diamonds represent absorption of the EGFP samples measured at λ=280 nm.*

Table S3 *Extinction coefficients of EGFP at different pH and wavelengths, determined using the calibration curves of figure 13.*

Buffer / λ	pH 5.0	pH 7.5
280 nm	26348	31807
395 nm	28506	
490 nm		55000[a]

[a] literature value[7]

Detection of GE using anti-EGFP antibodies

A sample containing capsids loaded with 7 EGFP proteins, was divided into two identical samples. To one of these samples anti-EGFP antibody conjugated with rhodamine (Santa Cruz biotechnology) dissolved in buffer pH 5.0 was added in an approximate 1:1 molar ratio with the encapsulated GE proteins, and to the other sample the same volume of buffer was added. The antibodies were allowed to bind for 1 h at room temperature, before analysis on an FPLC system equipped with a Superose 6 column. The same was done for sample containing empty capsid.

To test the binding of the anti-EGFP antibody to GE in buffer pH 5.0 a small drop of GE and empty capsid (as a negative control) in pH 5.0 buffer was added to a nitrocellulose membrane.

minutes. The solution was removed and anti-EGFP antibody in 1% (w/v) non-fat dry milk in buffer pH 5.0 was added to the membrane and allowed to bind for 1 h. Non-bound antibody was then washed away with buffer pH 5.0 and a secondary antibody (Rabbit anti-mouse HRP) in 1% (w/v) non-fat dry milk in TBST (0.01 M Tris-HCl, 0.15 M NaCl, 0.05% Tween-20, pH 8.0) was allowed to bind overnight at 4°C. Non-bound secondary antibody was washed away with TBST and the membrane was incubated with an enhanced chemiluminescence (ECL™) solution (GE healthcare life sciences) for 2 min. and subsequently exposed to film (Kodak medical X-ray film) for two min. before development.

Calculation of GE:CK ratio in GE-CK complex

The capsid protein has a peak absorption at λ=280 nm, and at pH 7.5 EGFP has two peak absorptions, a standard protein absorption at λ=280 nm and an EGFP specific absorption at λ=490 nm. The ratio of the absorbance at λ=280 nm and λ=490 nm of the GE-CK complex at pH 7.5 was determined by UV-Vis spectrophotometry (figure 4) and was found to be 1.07. According to the law of Lambert-Beer, absorption (A) is equal to the extinction coefficient (ε) times the concentration (c) times the path length of the cuvet (d). Since the path length is the same in all cases, we can omit this parameter from the equation, and substitute the absorption by the extinction coefficient times the concentration (the extinction coefficients of EGFP has been previously calculated, and the extinction coefficient of CK was assumed to be the same as that of the wild-type capsid. i.e. ε =24075). After rearrangement of the resulting formula, this results in a ratio of GE:CK of 0.89. Hence, approximately 0.9 GE proteins are bound to 1 capsid protein monomer.

$$\frac{A280}{A490} = \frac{A280_{CK} + A280_{GE}}{A490_{GE}}$$

$$A = \varepsilon c$$

$$\frac{A280}{A490} = \frac{\varepsilon 280_{CK} c_{CK} + \varepsilon 280_{GE} c_{GE}}{\varepsilon 490_{GE} c_{GE}}$$

$$\frac{A280}{A490} \varepsilon 490_{GE} c_{GE} = \varepsilon 280_{CK} c_{CK} + \varepsilon 280_{GE} c_{GE}$$

$$\frac{A280}{A490} \varepsilon 490_{GE} = \frac{\varepsilon 280_{CK} c_{CK}}{c_{GE}} + \varepsilon 280_{GE}$$

$$\frac{\varepsilon 280_{CK}}{\frac{A280}{A490} \varepsilon 490_{GE} - \varepsilon 280_{GE}} = \frac{c_{GE}}{c_{CK}}$$

Calculation of the GE-CK complex concentration

The GE-CK complex concentration was determined by UV-Vis spectrophotometry. The extinction coefficient of the complex was assumed to be the extinction coefficient of CK plus the extinction coefficient of GE times the ratio of GE:CK ($\varepsilon_{CK}+\varepsilon_{GE}\times 0.9$).

Calculations of EGFP encapsulation

The number of EGFP proteins in a capsid can be calculated similarly to the number of GE molecules bound per capsid protein, by using the absorption ratios of the capsid peak (V=1.1 mL) from the FPLC data.

The capsid protein has a peak absorption at λ=280 nm, and at pH 5.0 EGFP has two peak absorptions, a standard protein absorption at λ=280 nm and an EGFP specific absorption at λ=395 nm. Therefore, the ratios of these absorptions can be used to determine the ratio of CK: GE in a mixture. The GE-CK complex was purified by FPLC, and the filled capsid eluted at V=1.1 mL. The ratios of the areas under the peak at V=1.1 mL for λ=280 nm and λ=395 nm were determined. All proteins, wt CP, CK and GE, contribute to the absorption at λ=280 nm, but only GE contributes to the absorption at λ=395 nm. Hence, after substituting the absorption with the extinction coefficient times the concentration, and rearranging the formula, this results in the ratio of the concentration of all capsid proteins to EGFP proteins. Since there are 180 capsid proteins per capsid, the number of EGFP molecules per capsid can easily be derived from this ratio.

$$\frac{A280}{A395}=\frac{A280_{CP+CK}+A280_{GE}}{A395_{GE}}$$

$$A=\varepsilon c$$

$$\frac{A280}{A395}=\frac{\varepsilon 280_{CP+CK}c_{CP+CK}+\varepsilon 280_{GE}c_{GE}}{\varepsilon 395_{GE}c_{GE}}$$

$$\frac{A280}{A395}\varepsilon 395_{GE}c_{GE}=\varepsilon 280_{CP+CK}c_{CP+CK}+\varepsilon 280_{GE}c_{GE}$$

$$\frac{A280}{A395}\varepsilon 395_{GE}=\frac{\varepsilon 280_{CP+CK}c_{CP+CK}}{c_{GE}}+\varepsilon 280_{GE}$$

$$\frac{\frac{A280}{A395}\varepsilon 395_{GE}-\varepsilon 280_{GE}}{\varepsilon 280_{CP+CK}}=\frac{c_{CP+CK}}{c_{GE}}$$

$$\text{number of EGFP proteins per capsid}=\frac{c_{CP+CK}}{c_{GE}}\Big/180$$

The number of EGFP proteins encapsulated per capsid however does not indicate which percentage of the added EGFP is encapsulated into the capsids. To calculate this, the amount of EGFP proteins per capsid is divided by the maximum number of EGFP proteins per capsid (ie. the number of EGFP proteins that could theoretically have been encapsulated if all EGFP

proteins would have been incorporated into a capsid, and all capsid proteins would have formed a capsid). When plotted against the EGFP-capsid complex to total protein ratio (i.e. the percentage of EGFP added to the sample), the resulting graph shows an exponential decay in the percentage of EGFP which is incorporated into the capsids (figure 14).

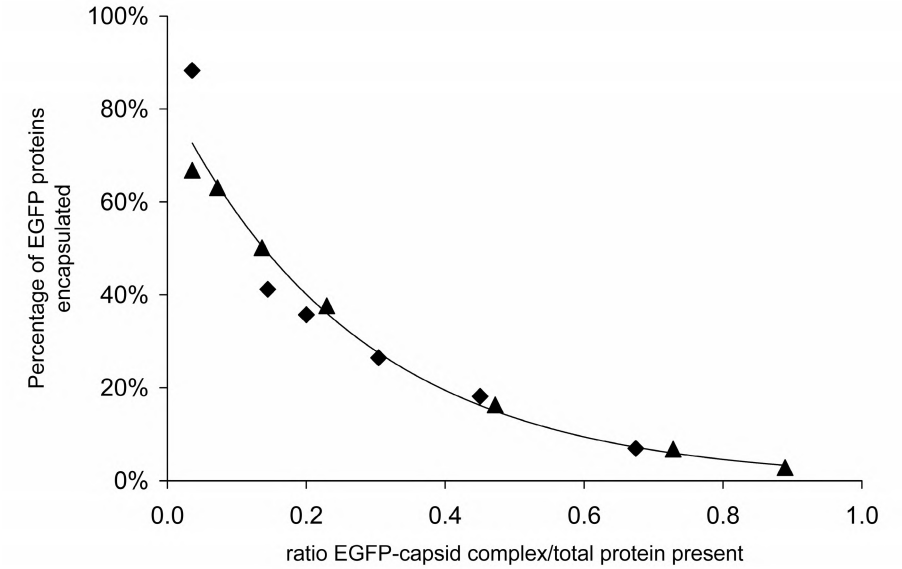

Figure 14 *Graph of the percentage of EGFP proteins encapsulated as a function of the GE-CK complex to total protein ratio. Diamonds and triangles represent data points of duplicate experiments. The line represents the exponential trend line though the data points.*

References

1. Douglas, T. and M. Young, *Host-guest encapsulation of materials by assembled virus protein cages.* Nature, **1998**, *393*, 152.
2. Loo, L., et al., *Infusion of dye molecules into Red clover necrotic mosaic virus.* Chem. Commun., **2008**, 88.
3. de la Escosura, A., et al., *Viral capsids as templates for the production of monodisperse Prussian blue nanoparticles.* Chem. Commun., **2008**, 1542.
4. Young, M., et al., *Plant viruses as biotemplates for materials and their use in nanotechnology.* Annu. Rev. Phytopathol., **2008**, *46*, 361.
5. Comellas-Aragones, M., et al., *A virus-based single-enzyme nanoreactor.* Nat. Nanotechnol., **2007**, *2*, 635.
6. Sikkema, F.D., et al., *Monodisperse polymer-virus hybrid nanoparticles.* Org. Biomol. Chem., **2007**, *5*, 54.
7. Litowski, J.R. and R.S. Hodges, *Designing heterodimeric two-stranded alpha-helical coiled-coils - Effects of hydrophobicity and alpha-helical propensity on protein folding, stability, and specificity.* J. Biol. Chem., **2002**, *277*, 37272.
8. Burkhard, P., J. Stetefeld, and S.V. Strelkov, *Coiled coils: a highly versatile protein folding motif.* Trends in Cell Biology, **2001**, *11*, 82.
9. Siegert, R., et al., *Structure of the molecular chaperone prefoldin: Unique interaction of multiple coiled coil tentacles with unfolded proteins.* Cell, **2000**, *103*, 621.
10. Tripet, B., et al., *Engineering a de novo-designed coiled-coil heterodimerization domain for the rapid detection, purification and characterization of recombinantly expressed peptides and proteins.* Protein Eng., **1996**, *9*, 1029.
11. Contarino, M.R., et al., *Modular, self-assembling peptide linkers for stable and regenerable carbon nanotube biosensor interfaces.* Journal of Molecular Recognition, **2006**, *19*, 363.
12. Wang, C., R.J. Stewart, and J. Kopecek, *Hybrid hydrogels assembled from synthetic polymers and coiled-coil protein domains.* Nature, **1999**, *397*, 417.
13. Robson Marsden, H., et al., *A Reduced SNARE Model for Membrane Fusion* Angew. Chem. Int. Edit., **2009**, *48*, 2330.
14. Dong, H. and J.D. Hartgerink, *Short homodimeric and heterodimeric coiled coils.* Biomacromolecules, **2006**, *7*, 691.
15. Bujarski, J.J., *Bromovirus Isolation and RNA Extraction.* Vol. 81. 1998. 183.
16. Verduin, B.J.M., *Degradation of Cowpea Chlorotic Mottle Virus Ribonucleic-Acid Insitu.* J. Gen. Virol., **1978**, *39*, 131.
17. Verduin, B.J.M., *Preparation of CCMV-Protein in Connection with Its Association into a Spherical-Particle.* FEBS Lett., **1974**, *45*, 50.
18. Baird, G.S., D.A. Zacharias, and R.Y. Tsien, *Biochemistry, mutagenesis, and oligomerization of DsRed, a red fluorescent protein from coral.* P. Natl. Acad. Sci. USA, **2000**, *97*, 11984.

.

Chapter 7

*Complex assembly behavior during the encapsulation of enhanced green fluorescent protein in the cowpea chlorotic mottle virus capsid**

Introduction

In the previous chapter we have presented a method for the efficient and controlled encapsulation of EGFP proteins into the CCMV virus capsid.[1] In short, this was accomplished via the non-covalent attachment of EGFP to the capsid proteins via a short heterodimeric coiled-coil sequence. EGFP was provided with the E-coil (GE) via molecular biology techniques, and the capsid protein was likewise provided with the K-coil (CK). Mixing of the two proteins resulted in an EGFP-capsid protein complex (GE-CK complex). In order to encapsulate the EGFP, this complex was mixed with wt CP in various ratios, to allow control over the amount of encapsulated EGFP and to prevent overcrowding of the capsid, which forms upon lowering the pH.

The formed capsids contained different amounts of EGFP depending on the initial [GE-CK complex]/[total protein] ratio. When this ratio increases the number of encapsulated EGFP proteins per capsid first increases and then decreases again. Clearly, there must be a maximum to the number of proteins that can be encapsulated, due to limited space inside the capsid. But there does not seem to be an obvious reason why the number of encapsulated EGFP proteins per capsid should decrease with increasing amounts of EGFP present. In this chapter we provide an explanation of this seeming discrepancy between theory and practice.

* This work was published: *Macromol Biosci*, **2010**, 10, 539

Results and Discussion

Capsids containing multiple encapsulated EGFP molecules per capsid were analyzed by FPLC (figure 7, chapter 6) EGFP has a specific absorption at λ=395 nm at pH 5.0. By monitoring both this EGFP specific absorption and the protein absorption at λ=280 nm during the run, EGFP can be distinguished from the capsid protein. Formed capsids typically elute at V=1.1 mL and since the EGFP specific absorption is present at V=1.1 mL, this indicates that EGFP is encapsulated. Free proteins like EGFP, and the capsid dimers elute between V=1.6-1.9 mL. A careful look at the FPLC graphs of the encapsulation experiments, shows that the peak around V = 1.6 – 1.9 shifts to lower elution volumes with increasing [GE-CK complex]/[total protein] ratios, which is indicative of the formation of a larger complex (figure 1).

It is hypothesized that this unknown complex (UC) consists of several GE-CK components. If this is true, the GE-CK contained in this UC would presumably be unavailable for incorporation into the capsid structure. This might explain the decrease in EGFP encapsulation efficiency with increasing [GE-CK complex]/[total protein] ratios.

To investigate this hypothesis, we first examined whether the UC was solely composed of the GE-CK complex, or of the GE-CK complex and the wt CP. This was done by labeling wt CP with a fluorescent dye (DyLight™ 649) and repeating an encapsulation experiment with a [GE-CK complex]/[total protein] ratio that was known to cause the formation of the UC. The sample was analyzed by FPLC equipped with a Superose 6 column, and the resulting graph again showed a peak associated with the UC at V=1.65 mL, but the labeled wt CP eluted at V=1.73 mL (Figure 2). Hence, the UC seems to be composed of the GE-CK complex.

In the encapsulation experiments, the concentration of GE-CK complex was increased, while keeping the total concentration of capsid proteins (wt CP and CK) approximately at an equal level, to obtain increasing [GE-CK complex]/[total protein] ratios. As a result the effect of the concentration of the GE-CK complex on the UC formation is studied. During the GE-CK complex purification on FPLC it was already observed that at pH 7.5 the UC was formed, so the GE-CK complex was studied in more detail at this pH using a Superdex 200 FPLC column. This column is better suited for the mass range of the complex and the UC than the Superose 6 column used for the encapsulation experiments.

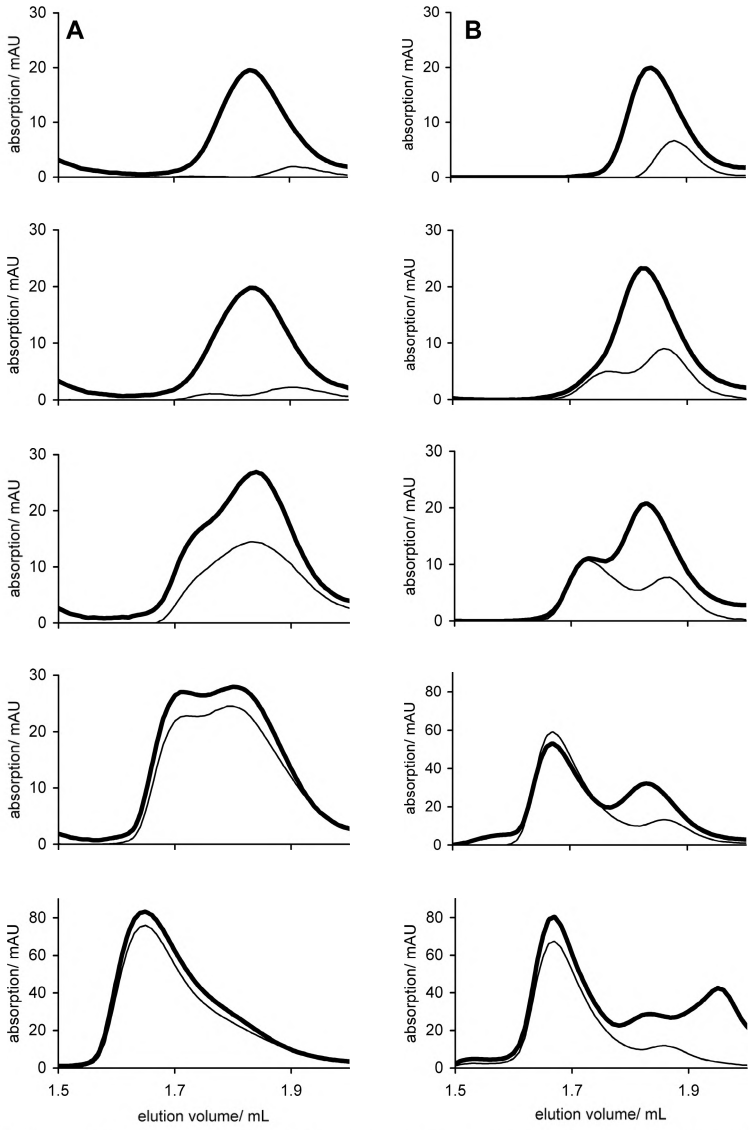

Figure 1 *Part of the FPLC graphs of the various encapsulation experiments. The FPLC was equipped with a superose 6 column. Thick lines represent the protein absorption at λ=280 nm and thin lines the EGFP absorption at λ=395 nm. A) Graphs of the encapsulation experiments with increasing amounts of GE-CK complex. [GE-CK complex]/[total protein] ratios increase in the following order from top to bottom: 0.03, 0.14, 0.29, 0.43 and 0.66. B) Graphs from encapsulation experiments in which uncomplexed GE was used. [GE]/[total protein] ratios increase in the following order from top to bottom: 0.03, 0.12, 0.20, 0.43 and 0.69.*

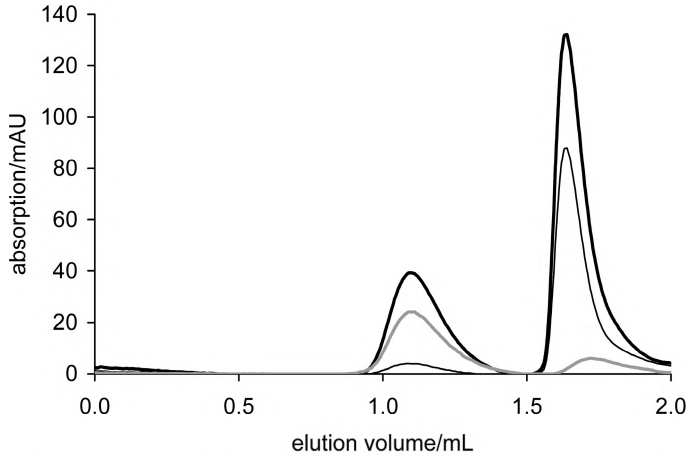

Figure 2 *FPLC graph of GE-CK complex mixed with labeled wt CP in a [GE-CK complex]/[total protein] ratio of 0.7 and dialyzed to pH 5.0. The FPLC was equipped with a Superose 6 column. The thick black line represents protein absorption at λ=280 nm, the thin black line represents EGFP absorption at λ=395 nm, and the grey line represents labeled wt CP absorption at λ=647nm.*

The complexes were studied at several concentrations spanning the range of the concentrations used for the encapsulation experiments. The elution volumes increased with increasing GE-CK complex concentrations, this is indicative of the formation of a larger complex, presumably the UC. In order to estimate the molecular weights corresponding to the elution volumes of these complexes, a calibration curve was made for this column (see chapter 6).

This calibration curve was then used to calculate the molecular weights of the formed complexes. The resulting graph of GE-CK complex concentration versus molecular weight (figure 3), shows a gradual increase in molecular weight with increasing GE-CK concentrations. Since the proteins of which the complex consists, have a discrete mass, a gradual increase can best be explained by a shifting equilibrium between two, or more distinct forms of the GE-CK complex. According to the calibration curve, the molecular weights corresponding to the smallest elution volume, V=1.3 mL and the largest elution volume, V=1.4 mL are 239 kDa and 127 kDa, respectively. The theoretical molecular weight for the GE-CK complex is 113 kDa, which corresponds to the V=1.4 mL peak. The molecular weight corresponding to the 1.3 mL peak (239 kDa) could correspond to two times the molecular weight of the GE-CK complex i.e. 226 kDa. Thus, the two distinct forms

that are in equilibrium could be the GE-CK complex and a complex of twice that molecular weight. Since the graph of GE-CK complex concentration versus MW does not level off completely at the highest concentration used for EGFP encapsulation, it is possible that at higher concentrations even bigger complexes are formed. This was not studied in further detail.

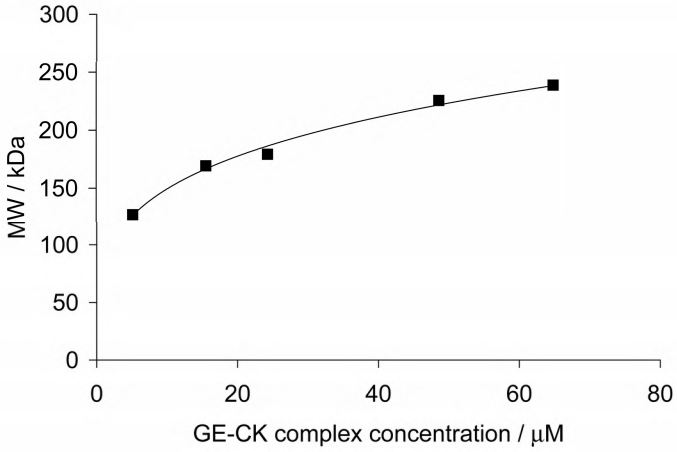

Figure 3 *Graph of GE-CK complex concentration versus the molecular weight of the resulting complex. Closed squares represent data points. The elution volumes corresponding to the calculated molecular weights of the data points are: 1.40 mL for 127 kDa, 1.35 mL for 169 kDa, 1.34 mL for 179 kDa, 1.30 mL for 225 kDa and 1.29 mL for 239 kDa.*

Since the capsid formation is induced by lowering the pH from 7.5 to 5.0, it was also investigated whether the same effect would be seen at pH 5.0. This indeed seemed to be the case. At low concentrations of the GE-CK complex (6.3 µM), the complex had an elution volume of V=1.4 mL, which corresponds to the molecular weight of the GE-CK complex. At high concentrations (37.9 µM), the elution volume had shifted to V=1.3 mL, corresponding to the molecular weight of two times the GE-CK complex.

To determine whether the elution volume of the UC on a Superose 6 FPLC column (V~1.6 mL) also corresponded to twice the molecular weight of the GE-CK complex, a calibration curve was also made for the Superose 6 column (see figure 5b, chapter 6). This indicated that the elution volumes of the GE-CK complex and that of a complex of twice that weight should have an elution volume of V=1.73 mL

and V=1.64 mL, respectively. This in turn suggests that the peak of the UC at V=1.6 mL on the Superose 6 column indeed originates from a complex that has twice the weight of the GE-CK complex.

Since it seems that the UC is twice the molecular weight of the GE-CK complex, the UC might be a dimeric form of the GE-CK complex. A possible explanation is that the EGFP proteins of the two GE-CK complexes dimerize. The crystal structure of GFP shows two GFP molecules dimerized in an antiparallel fashion,[2,3] and EGFP also has a weak tendency to dimerize.[4] GE also seems to have a tendency to dimerize at high concentrations, as was suggested by FPLC data, measured on an FPLC equipped with a Superdex 200 column. At low concentrations (16 μM) GE had an elution volume of V=1.64 mL on this column, which corresponds to a molecular weight of 32 kDa. The theoretical molecular weight for monomeric GE is 33 kDa. High concentrations of GE (390 μM) however, resulted in an elution volume of V=1.51 mL, which corresponds to a molecular weight of 68 kDa. This is about twice the mass of the monomeric GE, thus indicating dimerization of GE proteins and making it conceivable that this phenomenon is the cause of the apparent dimerization of the GE-CK complex. The GE proteins bound to the capsid protein are positioned parallel with respect to each other (figure 4), whereas the GFP dimers in the crystal structure are dimerized in an antiparallel fashion. It is possible that this parallel orientation prevents the internal dimerization of the GE proteins in the GE-CK complex. GE proteins of two different GE-CK complexes, however, might be able to align in an antiparallel fashion in that way enabling the formation of a dimeric complex. Hence, it is possible that the UC is a dimeric form of the GE-CK complex.

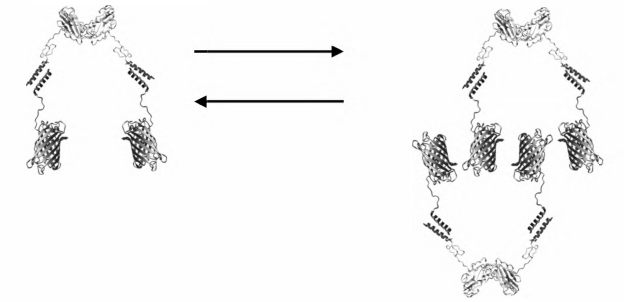

Figure 4 *Schematic representation of the structure of the CK-GE complex and the possible structure of the CK-GE complex dimer.*

Besides the concentration of the GE-CK complex, other factors might also have an influence on the formation of the dimeric GE-CK complex. Since in the encapsulation experiments not only the concentration of the GE-CK complex is changed, but also the [GE-CK complex]/[total protein] ratio, the effect of this ratio was also investigated. The proteins were mixed in a ratio that was known to cause the formation of the dimeric GE-CK complex, i.e. a ratio of 0.7 (figure 1), and dialyzed to buffer pH 5.0. This mixture was measured on an FPLC equipped with a Superose 6 column, and as expected, the majority of the GE-CK complex was in the dimerized form (figure 5a). A very concentrated solution of the wt CP at pH 7.5 was then added to the mixture, thereby shifting the ratio from 0.7 to 0.4. Due to the small volume of the wt CP solution this did not significantly affect the pH, or the total concentration of GE-CK complex in the sample. The mixture was equilibrated for one night and was then again measured using FPLC. The resulting FPLC graph (figure 5b) showed a decrease in the amount of the dimeric GE-CK complex formed, and also an increase in the number of encapsulated EGFP proteins per capsid. This experiment shows that the [GE-CK complex]/[total protein] ratio also affects the formation of the dimeric GE-CK complex, and that it is a dynamic equilibrium.

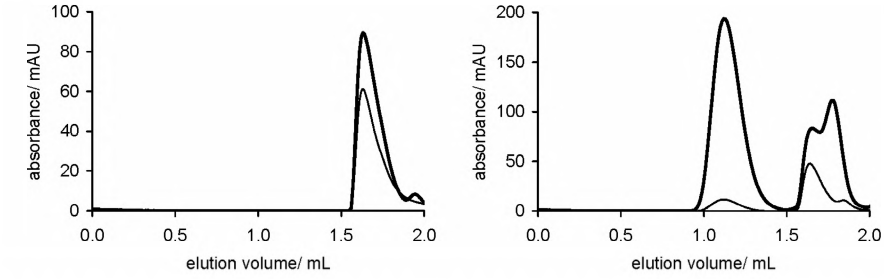

Figure 5 *FPLC graphs of experiments to investigate the effect of the [GE-CK complex]/[total protein] ratio on the ability of the complex to form capsids. The FPLC system was equipped with a Superose 6 column and the samples were measured at pH 5.0. Thick lines represent protein absorption at λ=280 nm and thin lines represent EGFP absorption at λ=395 nm. A) [GE-CK complex]/[total protein] ratio of the sample is 0.7. B) [GE-CK complex]/[total protein] ratio of the sample was changed from 0.7 to 0.4*

Since it appears that both the concentration of the GE-CK complex and the [GE-CK complex]/[total protein] ratio have an effect on the formation of the dimeric GE-CK complex, and both these variables were changed in the

encapsulation experiments, new encapsulation experiments were performed, in which the concentration of the GE-CK complex was kept constant (experiment #3 and #4, see experimental section). That is, only the [GE-CK complex]/[total protein] ratio was varied. This resulted in a more or less constant encapsulation of about 9 EGFP proteins per capsid (figure 6).

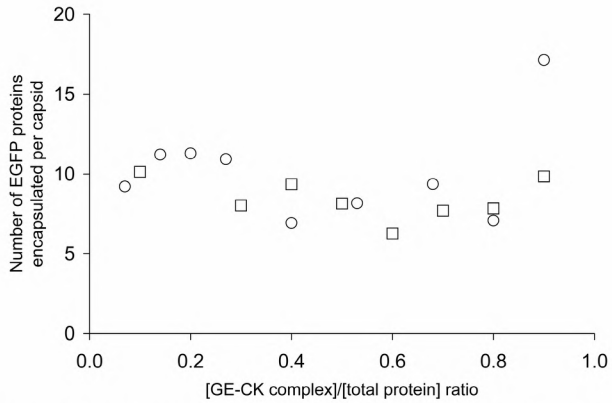

Figure 6 *Number of EGFP proteins encapsulated per capsid as function of the GE-CK complex/ total protein ratio. To obtain different ratios only the concentration of the wt CP was varied, the concentration of GE-CK was kept constant. Open squares depict data points of experiment #3 and open circles depict data points of experiment #4. Experiment #3 and #4 are duplicate experiments (see experimental section).*

The form of this graph is significantly different from that of the previous encapsulation experiments, in which the concentration of capsid proteins (wt CP and CK) instead of the concentration of the GE-CK complex was kept constant. This indicates that not only [GE-CK complex]/[total protein] ratio is of importance for the encapsulation efficiency of EGFP into the capsids, but that the absolute concentrations of the individual proteins are also critical.

It is likely that the dimeric GE-CK complex can not be incorporated into the capsid. This would explain the decrease in encapsulation efficiency with increasing [GE-CK complex]/[total protein] ratios. However, this does not appear to be the only cause, since the FPLC peaks around 1.7 mL that are probably caused by the dimeric GE-CK complex are also observed when negative control encapsulation experiments are performed in which only GE is added to the wt CP instead of the GE-CK complex. The peaks around V = 1.6 – 1.9 mL of the resulting FPLC graphs (figure 1b) showed a remarkable similarity to the peaks around V = 1.6 – 1.9 mL of

the normal encapsulation experiments (figure 1a). So these peaks could also be caused by the GE proteins without CK. To check if this was also the case for the normal encapsulation experiments with the GE-CK complex, FPLC fractions of both the capsid peak at V=1.1 mL and the peaks around V = 1.6 – 1.9 mL were analyzed by SDS-PAGE analysis and visualized by silver staining (figure 7).

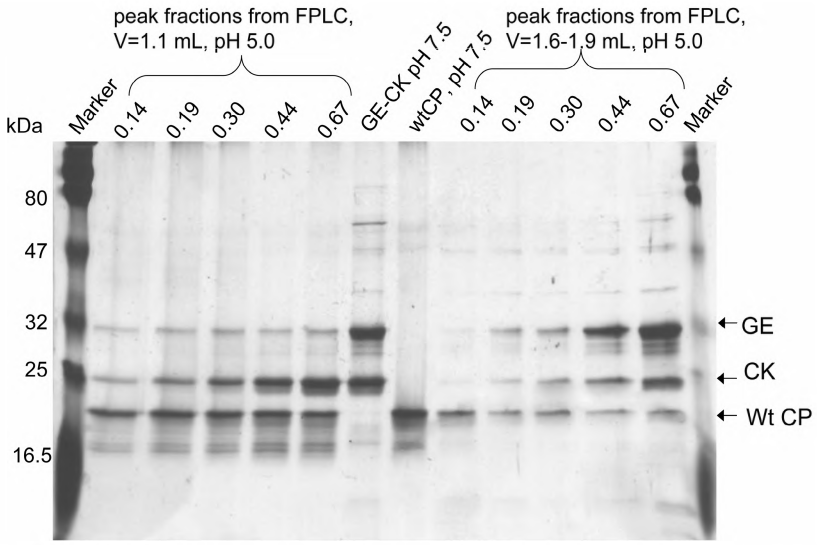

Figure 7 *Silver stained SDS-PAGE gel of EGFP encapsulation experiments with increasing amounts of GE-CK complex to increase the [GE-CK complex]/[total protein] ratio. Numbers above the lanes indicate this ratio. FPLC was equipped with a superose 6 column.*

This showed that the GE and CK proteins were not present in a 1:1 ratio in either the capsid peak at V=1.1 mL or the peaks between V=1.6 and 1.9 mL. This indicates that the GE-CK complex dissociates to certain extend, since otherwise the amount of GE and CK in the FPLC fractions would be equal. It is known from literature[5] that the E/K coil interaction is pH sensitive. At pH 7.5 the coiled-coil complex is stable, but at pH 5.0 the coiled-coils partially disassemble. Three disassembled E-coils can then form a trimeric coiled-coil structure. This would indicate that at pH 5.0 the (dimeric) GE-CK complex could partially disassemble and reassemble into free CK proteins and trimeric GE proteins. If this disassembly would take place faster than the capsid formation, this could influence the efficiency with which EGFP is encapsulated as it would result in a decrease of the amount of GE-CK complex in solution.

Conclusion

Inclusion of Enhanced Green Fluorescent Protein in the capsid of the Cowpea Chlorotic Mottle Virus can be controlled to a certain extent by introducing non-covalent interactions between the guest and the capsid proteins, resulting in the formation of a stable GE-CK complex at pH 7.5. The subsequent encapsulation process in the presence of wild type capsid displays an unexpected optimum when the [GE-CK complex]/[total protein] ratio is increased. It was expected that the GE-CK complex concentration in solution would be equal to the amount GE-CK complex added. This does not seem to be the case however. We have shown that at pH 7.5, part of the GE-CK complex tends to dimerize as a result of both increasing concentration of the complex as well as the increasing [GE-CK complex]/[total protein] ratio. This dimeric complex also seems to be present at pH 5.0. Presumably the dimeric complex cannot be incorporated into the capsid, so the actual GE-CK complex concentration in the mixture is lower than expected. Another factor contributing to the decrease in GE-CK complex concentration in the mixture, is the pH dependent dissociation of the GE-CK complex. At pH 5.0, the (dimeric) GE-CK complex partially dissociates into CK proteins and GE proteins. The latter will then probably re-associate to form a trimeric GE complex as is reported in detail for these coiled coil systems by Klok et al.[5]

So the mechanism of EGFP encapsulation as initially proposed by us (see chapter 6) and outlined in figure 6 of chapter 6 is more complex. Presumably part of the GE-CK complex dimerizes at pH 7.5 with increasing concentration and [GE-CK complex]/[total protein] ratio. At pH 5.0 this dimerized complex is not incorporated into the capsid. Due to the dissociation of the coiled-coils at this pH, both the dimerized GE-CK complex and the non-dimerized GE-CK complex probably partially disassemble, forming trimerized GE complexes and free CK. This leads to a mixture of five different species: the wt CP, free CK and GE-CK complex, which will assemble to a EGFP filled capsid, and the trimerized GE and dimerized GE-CK complex, which will not be incorporated into the capsid and elute from the FPLC column around V=1.7 mL (figure 8).

In summary, both the concentration and ratio dependent dimerization of the GE-CK complex as well as the pH dependent dissociation of the GE-CK complex result in a decrease of the available amount of GE-CK complex in solution, which in turn leads to a decrease in encapsulation efficiency.

Since formation of the dimeric GE-CK complex might be due to specific properties of the EGFP, it is possible that this effect will not occur with other proteins. This might mean that the current maximum of 15 EGFP proteins encapsulated per capsid is not necessarily the maximum for all proteins. Further studies are needed to answer this question.

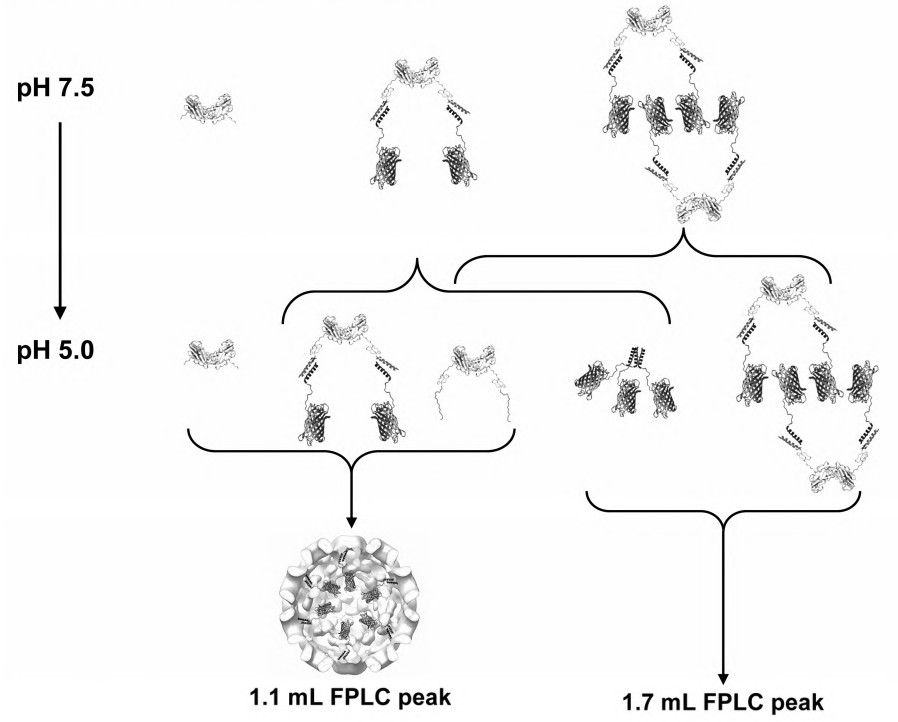

Figure 8 *Proposed EGFP encapsulation mechanism. The top row represents (from left to right): the wt CP, the GE-CK complex and the dimeric GE-CK complex at pH 7.5. Upon lowering the pH to 5.0 the GE-CK complex is thought to partially dissociate, resulting in a mixture of CK, trimeric GE and intact GE-CK complex. The dimeric GE-CK complex probably also partially dissociates, resulting in a mixture of CK, trimeric GE and dimeric GE-CK complex. The wt CP, together with GE-CK complex, and dissociated CK then presumably form a capsid with EGFP encapsulated, giving rise to a V=1.1 mL peak on the FPLC, while the trimeric GE and dimeric GE-CK complex probably give rise to the peaks between V=1.6 and 1.9 mL.*

Experimental section

Methods and materials

Protein concentrations were determined using a Cary 50 Conc (Varian, Middelburg) UV-VIS spectrophotometer using a quartz cuvette with a path length of 3 mm.

FPLC measurements were performed using either a Superose 6 PC 3.2/30 analytical column or a Superdex 200 PC 3.2/30 analytical column from GE lifesciences, on an Amersham Ettan LC system, fitted with a fractionating device. Buffers for FPLC were filtered with a Millipore 0.2 μM filter before use.

Two buffers were used: a buffer pH 7.5 (0.5 M NaCl, 0.05 M Tris-HCl, 0.01 M $MgCl_2$ and 0.001 M EDTA) and a buffer pH 5.0 (0.5 M NaCl, 0.05 M $NaCH_3COO$, 0.01 M $MgCl_2$ and 0.001 M EDTA). The pH of the buffers was set with HCl.

Expression and purification of wt CP, GE and CK

Expression and purification of wt CP, GE and CK, as well as the assembly of the GE-CK complex were carried out according to the procedures described in chapter 6.

EGFP encapsulation experiments

CK-GE complex or GE (negative control) in buffer pH 7.5 (50 mM Tris-HCl, 500 mM NaCl, 10 mM $MgCl_2$, 1 mM EDTA pH 7.5) were added in different ratios to wild type capsid protein in the same buffer (table 1 and table 2). The proteins were allowed to mix for 5 min., before dialyzing the mixture overnight to buffer pH 5.0. The mixtures were then analyzed on an FPLC system equipped with a Superose 6 column.

Encapsulation experiment with labeled wt CP

DyLight™ 649 NHS ester (Thermo Scientific) was added to wt CP in a 50 mM phosphate buffer pH 7.5 with 0.5 M NaCl, The unreacted dye was removed by extensive dialysis against buffer pH 7.5 and the labeling efficiency was determined according to the procedure described in the DyLight™ manual, which resulted in a labeling efficiency of 0.14 mol of the dye per mol wt CP. The labeled wt CP was mixed with GE-CK complex in a [GE-CK complex]/[total protein] ratio of 0.7. The concentration of wt CP in the sample was 2.01×10^{-5} M and that of GE-CK complex was 4.69×10^{-5} M. The mixture was then dialyzed to buffer pH 5.0 and analyzed by FPLC equipped with a Superose 6 column.

Table 1 *Data of the encapsulation experiments.*

Exp. #	M wt CP in samples	M GE-CK complex in samples	[a]Ratio [GE-CK complex]/ [total protein]	Number of EGFP proteins encapsulated per capsid
1	7.71×10^{-5}	2.50×10^{-6}	0.03	2.0
	7.39×10^{-5}	4.99×10^{-6}	0.06	2.4
	6.83×10^{-5}	9.36×10^{-6}	0.12	11.8
	6.02×10^{-5}	1.56×10^{-5}	0.21	14.8
	4.02×10^{-5}	3.12×10^{-5}	0.44	12.5
	2.01×10^{-5}	4.68×10^{-5}	0.70	7.3
	8.03×10^{-6}	5.62×10^{-5}	0.87	1.1
	0	6.24×10^{-5}	1.00	0.0
2	7.84×10^{-5}	2.52×10^{-6}	0.03	5.6
	6.86×10^{-5}	1.01×10^{-5}	0.13	10.3
	6.37×10^{-5}	1.39×10^{-5}	0.18	12.3
	5.47×10^{-5}	2.08×10^{-5}	0.28	13.7
	4.25×10^{-5}	3.03×10^{-5}	0.42	14.4
	2.45×10^{-5}	4.41×10^{-5}	0.64	7.7
3	4.72×10^{-5}	2.02×10^{-5}	0.30	8.0
	3.04×10^{-5}	2.02×10^{-5}	0.40	9.3
	2.02×10^{-5}	2.02×10^{-5}	0.50	8.1
	1.35×10^{-5}	2.02×10^{-5}	0.60	6.3
	8.68×10^{-6}	2.02×10^{-5}	0.70	7.7
	5.06×10^{-6}	2.02×10^{-5}	0.80	7.8
	2.25×10^{-6}	2.02×10^{-5}	0.90	9.8
4	2.69×10^{-4}	2.02×10^{-5}	0.07	9.2
	1.24×10^{-4}	2.02×10^{-5}	0.14	11.2
	8.10×10^{-5}	2.02×10^{-5}	0.20	11.3
	5.47×10^{-5}	2.02×10^{-5}	0.27	10.9
	3.04×10^{-5}	2.02×10^{-5}	0.40	6.9
	1.80×10^{-5}	2.02×10^{-5}	0.53	8.2
	9.53×10^{-6}	2.02×10^{-5}	0.68	9.4
	5.06×10^{-6}	2.02×10^{-5}	0.80	7.1
	2.25×10^{-6}	2.02×10^{-5}	0.90	17.1

[a)]Ratio was calculated by dividing molarity of the GE-CK complex by the molarity of the GE-CK complex+M wt CP.

Table 2 *Data of the negative controls for the encapsulation experiments. In these experiments uncomplexed GE was used instead of the GE-CK complex.*

Exp. #	M wt CP in samples	M GE in samples	[a]Ratio [GE]/ [total protein]	Number of EGFP proteins encapsulated per capsid
5	8.03×10^{-5}	0	0.00	1.00
	7.71×10^{-5}	2.38×10^{-6}	0.03	1.93
	6.83×10^{-5}	8.91×10^{-6}	0.12	1.84
	6.02×10^{-5}	1.48×10^{-5}	0.20	2.19
	4.02×10^{-5}	2.97×10^{-5}	0.43	2.20
	2.01×10^{-5}	4.45×10^{-5}	0.69	2.42
	8.03×10^{-6}	5.35×10^{-5}	0.87	2.75

Investigation of the [GE-CK complex]/[total protein] ratio on the aggregation behavior of GE-CK

GE-CK (8.28×10^{-5} M) was mixed with wt CP (9.55×10^{-5}M) in a GE-CK/total protein ratio of 0.7. The mixture was divided in two equal portions which were each dialyzed to buffer pH 5.0. After 2 hrs, a concentrated solution of wt CP (6.92×10^{-4}M, pH 7.5) was added to one of the mixtures resulting in a GE-CK/total protein ratio of 0.4. The two mixtures were further dialyzed to buffer pH 5.0 for approximately 10 h before analysis on an FPLC equipped with a Superose 6 column.

References

1. Minten, I.J., et al., *Controlled Encapsulation of Multiple Proteins in Virus Capsids.* J. Am. Chem. Soc., **2009**, *131*, 17771.
2. Tsien, R.Y., *The green fluorescent protein.* Annu. Rev. Biochem., **1998**, *67*, 509.
3. Yang, F., L.G. Moss, and G.N. Phillips, *The molecular structure of green fluorescent protein.* Nat. Biotechnol., **1996**, *14*, 1246.
4. Shaner, N.C., P.A. Steinbach, and R.Y. Tsien, *A guide to choosing fluorescent proteins.* Nat. Methods, **2005**, *2*, 905.
5. Apostolovic, B. and H.A. Klok, *pH-Sensitivity of the E3/K3 Heterodimeric Coiled Coil.* Biomacromolecules, **2008**, *9*, 3173.

Chapter 8

*Metal ion induced formation and stabilization of the cowpea chlorotic mottle virus capsid**

Introduction

The well-defined assemblies formed by virus capsid proteins turn out to be suitable construction materials for the formation of (functional) nanostructures.[1] The number of examples in which spherical viral cages are used as reaction vessels is steadily increasing,[2-4] Rod-like virus systems, on the other hand, are used as a scaffold for enzymes[5] or chromophores,[6] and as templates for the crystallization of solid state materials.[7] One of the next key steps in the further development of the emerging field of what is called "chemical virology", i.e. the use of virus based materials to direct or study chemical processes, is to further control the size, shape and stability of the well-organized architectures formed by virus capsid proteins. We are employing the CCMV capsid as such a well-defined scaffold, because of its pH dependent assembly behavior. From previous results (e.g. chapters 3 and 4 of this thesis) we know that subtle interactions with guest (macro)molecules can direct the assembly of the CCMV coat protein and that the charged N-terminus of the protein chain can play a role in this. As mentioned in chapter 5, the capsid protein was genetically modified with an N-terminal histidine tag to aid its purification. The N-termini of the capsid proteins are located at the quasi six-fold and five-fold axis of the capsid. Residues 29 to 33 of the N-termini of a protein hexamer located at the quasi-six fold axis form a hexameric tubular structure[8] (figure 1).

* This work is submitted for publication.

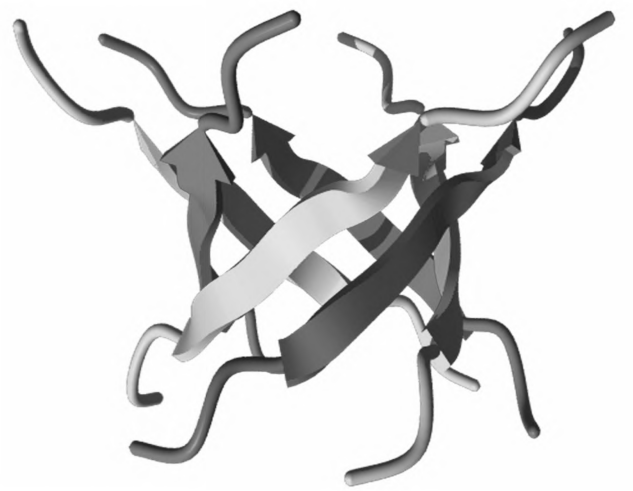

Figure 1 *Ribbon diagram showing the hexameric tubular structure formed by the N-termini (residues 26-39) of a CCMV hexamer located at the quasi-six fold axis.*

Like the N-termini of the coat proteins forming the hexamers, the five N-termini of the coat proteins forming the pentamers are protruding into the capsid at the center of the pentamers. The secondary structure of these N-termini appears to be unordered.

Histidines can coordinate certain metal ions, e.g. cobalt and nickel, and by immobilizing the metal ions on a resin functionalized with metal chelating groups like NTA, His-tag modified proteins can be selectively bound to the resin. Since the modified capsid protein is purified using Ni-NTA affinity chromatography, all purified capsid proteins are capable of coordinating to nickel atoms.

In this chapter we report that upon addition of nickel or other metal ions to the purified His-CPs, these ions appear to be able to induce and stabilize capsid assembly at neutral pH. Since the metal ions are positively charged and much smaller than negatively charged polyelectrolytes, the induction and stabilization of capsid assembly by metal ions is probably based on a different mechanism. The unprecedented stabilization of the T=3 capsid structure at pH 7.5 as described in this chapter makes the capsid more compatible with physiological conditions and hence a more applicable nano building block or nanoreactor. In this chapter we will describe in detail the induction of capsid-like structures by adding metal ions to the His-CP's dimers, as well as the stabilization of already assembled capsids from His-CP's by metal ions at neutral pH.

Results and Discussion

.

Induction of capsid like structures at pH 7.5 by nickel ions

The interaction of the His-tag sequence with metal salts is frequently used to immobilize proteins on a Ni-NTA colomn. The His-tag metal ion interaction can also be used to bind two or more His-tags together.[9] Since the modified CCMV coat proteins contain an N-terminal His-tag, it would be interesting to see if addition of nickel salts to the proteins would merely result in the connection between the two His-tags of a dimer, or that more complex structures are formed. Experimentally this was tested by the addition of 10 equivalents of $NiCl_2$ (i.e. 10 nickel ions per capsid protein monomer) to purified His-CP dimers in a phosphate buffer pH 7.5. The mixture was gently stirred at room temperature for one hour and analyzed using FPLC. A small fraction of the sample eluted at V=1.20 mL (figure 2), which indicates a particle which is significantly larger than a capsid protein dimer, and even larger than a T=1 particle (elution volume V=1.3 mL) but smaller than a T=3 particle (elution volume V=1.1 mL).

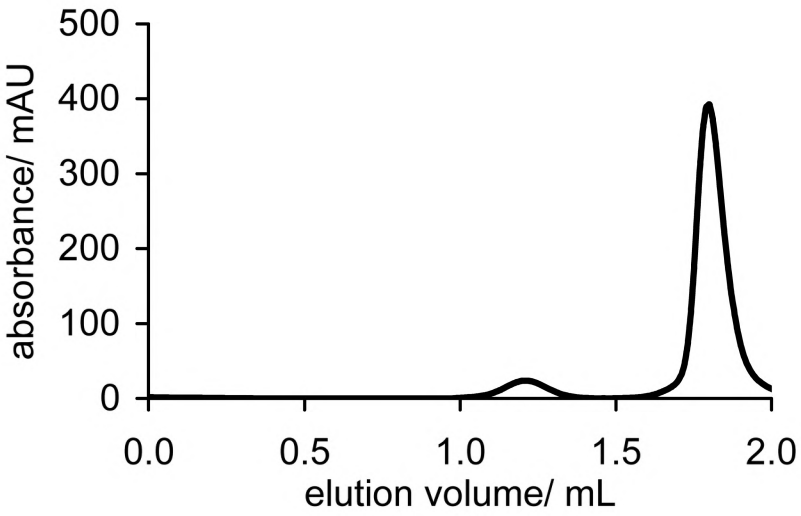

Figure 2 *FPLC trace of His-CP incubated with $NiCl_2$ for 1 hour at pH 7.5 The FPLC system is equipped with a Superose 6 column. The black line represents the protein absorption at λ=280 nm. The formed 22-24 nm particles elute at V=1.2 mL, and the unassembled His-CP dimers elute at ~1.8 mL.*

However, the majority of the proteins eluted at V=1.80 mL, which is the elution volume of the capsid protein dimer. To exclude that this could be caused by incorrect folding or other structural defect of the His-CP which might hinder the assembly of the capsid, the His-CP was assembled by lowering the pH to 5.0. The formed capsid eluted at V=1.1 mL, which is consistent with the elution volume of the wild-type capsid, and only a small portion of the His-CP's eluted at V=1.8 mL, confirming the ability of the His-CP's to form capsids. To investigate whether the formed particles were in a dynamic equilibrium with the capsid protein dimers, the fraction eluting at V=1.2 mL containing the formed particles was isolated and reinjected after 1 hour. The resulting FPLC graphs (figure 3) showed that almost all particles were still intact, indicating that the formed particles are stable. Based on the available data it was unclear if the material eluting at V = 1.8 mL contained bound nickel ions.

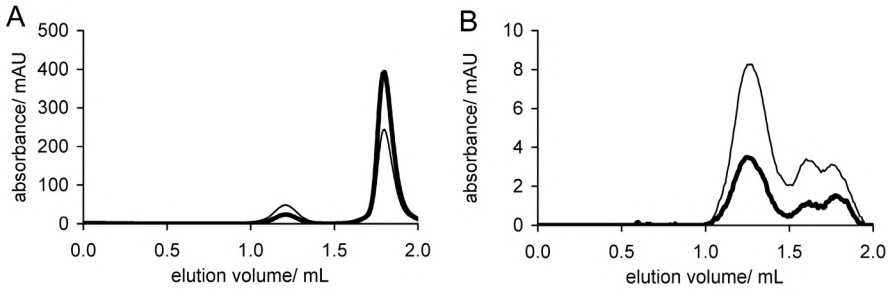

Figure 3 *A) FPLC trace of His-CP incubated with $NiCl_2$ for 1 hour at pH 7.5 The FPLC system is equipped with a Superose 6 column. The thick line represents the protein absorption at λ=280 nm, and the thin line the absorption at λ=260 nm. The formed 22-24 nm particles elute at V=1.2 mL, and the unassembled His-CP dimers elute at ~1.8 mL. B) Reinjection of collected fractions at 1.2 mL from graph A.*

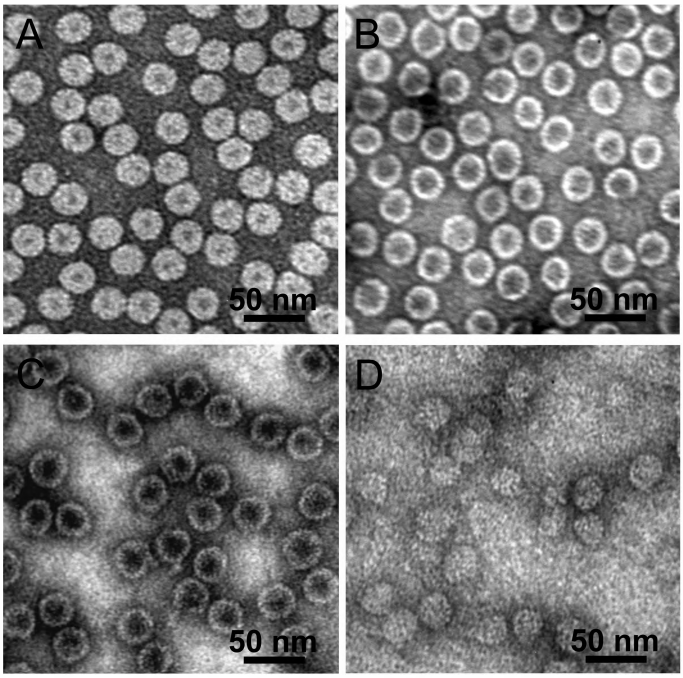

Figure 4 *TEM micrographs of uranyl acetate stained capsids. A) CCMV pH 5.0 B) Empty wild-type capsids pH 5.0 C) Empty His-CP capsids pH 5.0 D) Nickel induced His-CP particle formation pH 7.5*

To determine the size of the formed structures eluting at V=1.2 mL with greater precision, this fraction was analyzed by TEM which revealed the formation of capsid-like structures with an average diameter of D = 22 nm (figure 4). Unlike empty T=3 capsids, these capsids do not appear to have a stained interior. A stained interior of empty capsids is generally caused by the accumulation of the negative staining agent, uranyl acetate in this case, into the empty cavity. The importance of nickel atoms for the induction of the His-CP assembly was investigated by removing the nickel ions from the formed particles by adding two equivalents of EDTA with respect to nickel to the solution. No particles could be detected on TEM hereafter, indicating that the nickel ions are indeed essential for the formation and structural integrity of the particles. The necessity of the His-tag was studied separately, as other sites on the capsids are also known to have a strong affinity for metal ions. In particularly negatively charged residues located at the three-fold axis are known to have an affinity for calcium and other metal ions.[10] Binding of

calcium to these putative calcium binding sites is thought to be important in the stabilization of the virus at pH 5.0.[11] Given the affinity of these residues for positively charged ions, it is not unlikely that they will also coordinate to nickel ions. To investigate if coordination of nickel ions to the putative calcium binding sites might be responsible for the induction of particle formation, $NiCl_2$ was added to wt CP using the same procedure as described for the His-CP. Wt CP and His-CP both contain the putative calcium binding sites, but the wt CP lacks the His-tags, therefore any particle under these circumstances could not be attributed to nickel His-tag interaction. The results of these experiments revealed that no particles were formed from wt CP in the presence of nickel ions at pH 7.5, as was shown by TEM and FPLC analysis (figure 5). Hence both the His-tag and the nickel ions seem to be necessary for the assembly of the capsid-like structures.

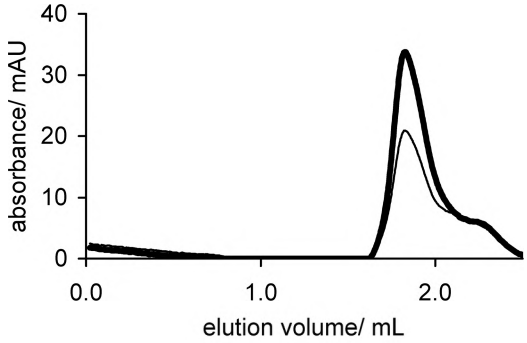

Figure 5 *FPLC trace of wt CP incubated with $NiCl_2$ for 1 hour at pH 7.5 The FPLC system is equipped with a Superose 6 column. The thick black line represents the protein absorption at λ=280 nm, and the thin black line the absorption at λ=260 nm. The protein dimers elute at V= 1.8 mL.*

Induction of capsid like structures at pH 7.5 by other metal ions

Nickel is a widely used metal-ion for the affinity purification of His-tagged proteins, but it is not the only metal ion that is capable of binding to histidines. Therefore, a range of other metal ions were tested for their ability to induce formation of the capsid-like structures. The procedure followed for the formation of these particles was identical to the procedure described for the $NiCl_2$ induced assembly, only the type of metal ion was varied. The results were analyzed by TEM (figure 4d and 6) and the following metals-salts were shown to induce particle formation: $RhCl_3$, $CoCl_2$, $PdCl_2$, and $AuCl_3$. Addition of $IrCl_3$, $PtCl_2$, $CuCl_2$, $ZnCl_2$, $AlCl_3$ and $AgNO_3$

to the His-CPs resulted in some capsid-like structure formation, but notably less compared to the previously mentioned metal ions and $NiCl_2$. No capsid-like structures were observed upon addition of $FeCl_2$ to the His-CPs. At this point the exact structure and the mechanism of formation of the His-tag containing proteins architectures in the presence of a variety of metal is still unclear. Since the absence of staining agent in the interior suggests either a solid particle or limited diffusion across the protein shell, the formed architectures are unlikely candidates for nanoreactors[12] and we, therefore, focused on a slightly different strategy for the use of metal ions in the capsid protein assembly process (see below).

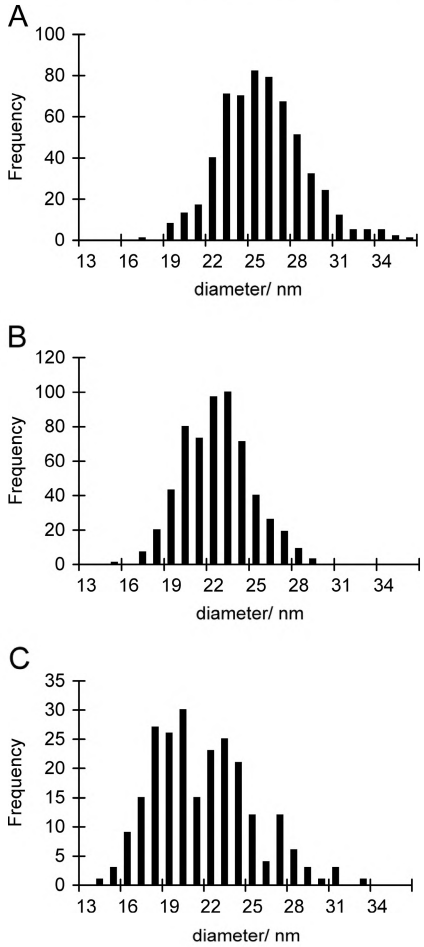

Figure 6 *Size distribution of His-CP particle formation induced at pH 7.5, based on TEM micrographs. A) Particle formation induced with $NiCl_2$. B) Particle formation induced with $RhCl_2$. C) Particle formation induced with $CoCl_2$.*

Stabilization of T=3 capsid structure at pH 7.5 by nickel ions

Although it was demonstrated that addition of Ni ions stabilized capsid-like structures at physiological pH it was clear from figure 3 that only a small part of the His-CPs was involved in this. This is in contrast with the pH induced assembly of His-CP into the native T=3 capsid structure, which does proceed with very high efficiency and yield. In this case, the formed capsids are only stable at acidic pH, which is often a drawback. Therefore we tried to induce the assembly of the capsid at pH 5.0, and to subsequently stabilize the formed assembly by the addition of $NiCl_2$ (10 eq.), after which the pH was increased to 7.5. The formed particles were analyzed after three days using FPLC, TEM and DLS. TEM analysis (figure 7a), showed particles of approximately 28 nm in size in line with FPLC results (figure 7b), where V = 1.15 mL points to the same size.

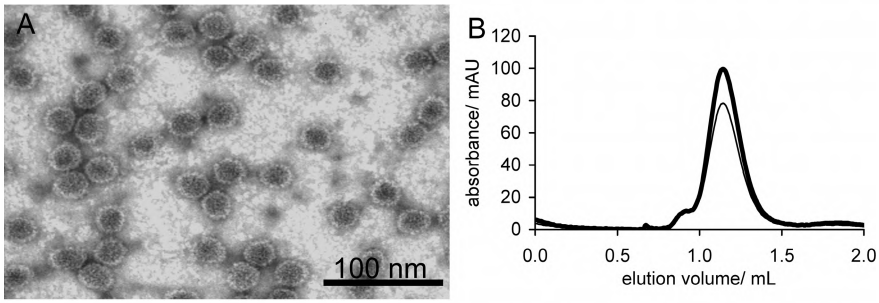

Figure 7 *TEM and FPLC results of nickel stabilized capsids after three days at pH 7.5, A) TEM micrographs of uranyl acetate stained nickel stabilized capsids. B) FPLC trace of nickel stabilized capsids at pH 7.5. The thick line represents the protein absorption at λ=280 nm, and the thin line the absorption at λ=260 nm. The formed capsids elute at V=1.15 mL. The FPLC system is equipped with a Superose 6 column.*

The capsids seem to be very mono-disperse and show a darkly stained interior, which is also observed with the wild-type capsids. The DLS volume distribution showed particles with a hydrodynamic radius of 30.4 nm (figure 8). This points to T=3 particles, which are assembled with high efficiency and these are stable at pH 7.5 for at least three days. To investigate the importance of nickel ions for the stabilization of the capsids at pH 7.5, EDTA was added. FPLC analysis of these samples now revealed an elution volume of V=1.8 mL, indicating that all capsids had disassembled into dimers (figure 9).

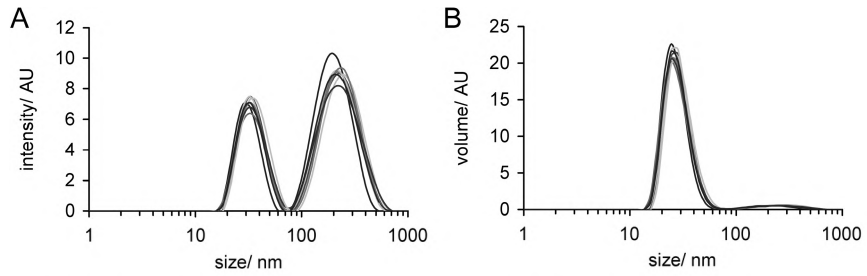

Figure 8 *DLS traces of nickel stabilized capsids at pH 7.5. Different lines indicate repeated measurements. A) DLS traces plotted against intensity. B) DLS traces plotted against volume.*

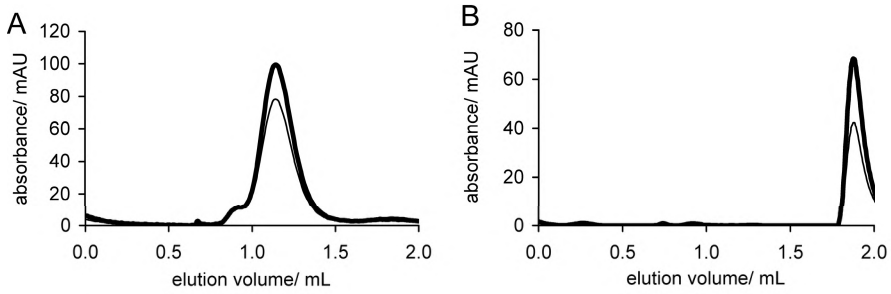

Figure 9 *FPLC traces showing the effect of addition of EDTA to the stabilized capsids. The thick lines represent the protein absorption at λ=280 nm, and the thin lines the absorption at λ=260 nm. The formed capsids elute at V=1.15 mL. Capsid protein dimers elute at V=1.8 mL. The FPLC system is equipped with a Superose 6 column. A) FPLC trace of nickel stabilized capsids at pH 7.5. B) FPLC trace of the same sample after addition of EDTA.*

No capsids could be observed by TEM analysis either, indicating that the nickel ions are indeed essential for the stabilization of the capsid at pH 7.5. To demonstrate that the stabilization of the capsid is caused specifically by the interaction of the His-tag with the nickel ions and not by interactions of other nickel-coordinating amino-acids, like the negatively charged residues at the putative calcium binding site, wt CP's, and His-CP capsids were separately labeled with chromophores and assembled by lowering the pH to 5.0. The formed capsids were then mixed and $NiCl_2$ was added before increasing the pH to 7.5 (figure 10a). The mixture was analyzed by FPLC; the His-CPs eluted at V=1.1 mL, and the wild-type capsid proteins eluted at V=1.8 mL. This clearly shows that the His-CP's did not disassemble upon increasing the pH, whereas the capsids of the wt CP did (figure

10b), confirming the need of the His-tag for the stabilization of the T=3 particle at pH 7.5. Since the two dyes used to label the capsid proteins have a small overlap at the measured wavelengths (figure 11) The small peak absorption at λ=646 nm can be completely attributed to assembled His-CP (see legend to figure 11).

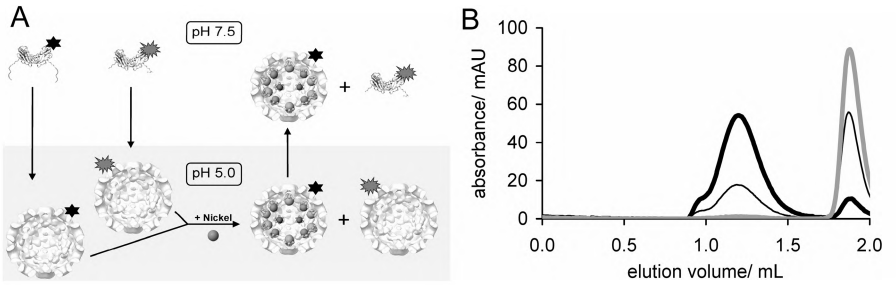

Figure 10 *Selective stabilization of His-CP capsids. A) Schematic representation of the experiment. Wt CP and His-CP were labeled with different dyes and the pH was subsequently lowered, resulting in wt CP capsid labeled with Dylight™ 647 and His-CP capsids labeled with Dylight™ 549. The capsids were then mixed and after stabilization of the capsids with nickel ions, the pH was increased back to pH 7.5. The resulting mixture was analyzed by FPLC. B) FPLC trace of the mixture at pH 7.5 after the stabilization step. The thin black line represents the protein absorption at λ=280 nm, the thick black line represents the absorption of the labeled His-CP at λ=562 nm and the grey line represents the absorption of the labeled wt CP at λ=646 nm. The stabilized capsids elute at V=1.15 mL and the capsid protein dimers elute at V=1.8 mL. The FPLC system is equipped with a Superose 6 column.*

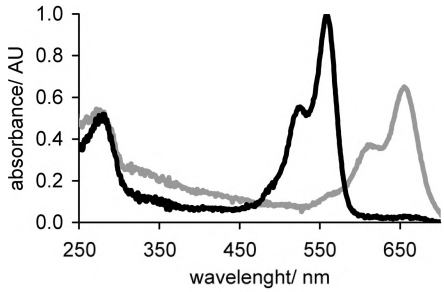

Figure 11 *UV-Vis traces of His-CP labeled with Dylight™ 549 (black line) and wt CP labeled with Dylight™ 647 (grey line). Labeled His-CP has a major absorption peak at λ= 562 nm, and also absorbs slightly at the wavelength of the major absorption peak of wt CP (λ= 646 nm). The absorption of labeled His-CP at λ= 562 nm is 48 times larger than the absorption of labeled His-CP at λ= 646 nm. In the FPLC graph shown in figure 10b a small absorption at λ= 646 nm is visible at V=1.15 mL. This absorption can be completely attributed to absorption of His-CP at λ= 646 nm. The area of the V=1.15 mL peak at λ= 562 nm is 16.31, and the area of the V=1.15 mL peak at λ= 646 nm is 0.29. The expected area at λ= 646 nm would be 16.31/48= 0.34. The expected and observed areas are very similar, and it is thus very likely that the small absorption at λ= 646 nm at V=1.15 mL is indeed caused by the presence of labeled His-CP and not due to assembled wt CP. The slight difference between the experimental value of 0.29 and the theoretical value of 0.34 can be attributed to inaccuracies in the measurement of these small values.*

The concentration of nickel ions bound to the capsid was determined by inductively coupled plasma-optical emission spectroscopy (ICP-OES). A sample of stabilized capsids was purified by FPLC and the concentration of sulfur and nickel in the fraction that eluted at V = ~1.15 mL was determined with ICP-OES (table 1). The concentration of nickel was just above the detection limit of ICP-OES and was therefore confirmed by inductively coupled plasma-mass spectrometry (ICP-MS), a more sensitive technique for the detection of nickel. Sulfur cannot be analyzed by ICP-MS. The number of nickel atoms bound per protein monomer could be derived from the Ni/S ratio measured using ICP-OES, as the amount of sulfur containing amino-acids in His-CP is known (two cysteines and three methionines, the starting methionine is removed by *E.coli* during expression), resulting in 0.7 nickel atoms per His-CP monomer. At pH 5.0, the theoretical number of nitrogen atoms of the imidazole side-chains that are non-protonated and thus able to coordinate to nickel atoms is 0.6 per His-tag. This number suggests that at pH 5.0 the available

imidazole groups are bound to nickel in a one to one ratio. At pH 7.5 this ratio will be higher because of the larger number of available imidazole groups. This is expected to lead to cross-linking and stabilization of the subunits.

Table 1 *Concentration of nickel and sulfur in stabilized T=3 particles at pH 7.5 determined by ICP-OES and ICP-MS*

	Concentration Ni (ppb)	Concentration S (ppb)
ICP-OES	97.75	676.1
ICP-MS	94.78	-

Formation of capsids containing wt CP's and His-CP's at pH 7.5

Both the presence of the His-tag in the His-CP as the addition of nickel ions to the sample seem to be essential for the stabilization of the T=3 capsid at pH 7.5. It is, however, not known if all 180 His-tags present in the capsid are needed for capsid stabilization at pH 7.5, or if a small amount of capsid protein lacking the His-tag would be tolerated. To test this, the amount of His-CP in the capsids was varied by mixing wt CP and His-CP in various ratios. Gillitzer et al.[4] already showed that differently labeled dimers can be incorporated into the same CCMV capsid. Since the wt CP and His-CP dimers are structurally very similar, presumably they would also form mixed capsids at pH 5.0. To be able to differentiate between the wt CP and the His-CP they were labeled with different dyes. The labeled protein dimers were subsequently mixed in different ratios at pH 7.5 and the mixture was assembled at pH 5.0. $NiCl_2$ was then added and the pH was increased to 7.5 (figure 12a).

The mixtures were analyzed by FPLC (figure 12b), and the amount of assembled capsid was derived from the ratio between the V=1.1 mL and V=1.8 mL peaks. For this the absorption at λ=280 nm for the peak at V=1.1 mL was divided by the total absorption at λ=280 nm, i.e. the amount of formed capsid was divided by the total amount of protein present in the sample, giving the percentage of assembled capsid. When more wt CP was present in the mixture, the amount of assembled capsid decreased (figure 13a). At low wt CP / total protein ratios the amount of assembled capsid protein was higher than would be expected if only the His-CPs would be able to assemble, indicating that at least part of the wt CP is also involved in the assembly.

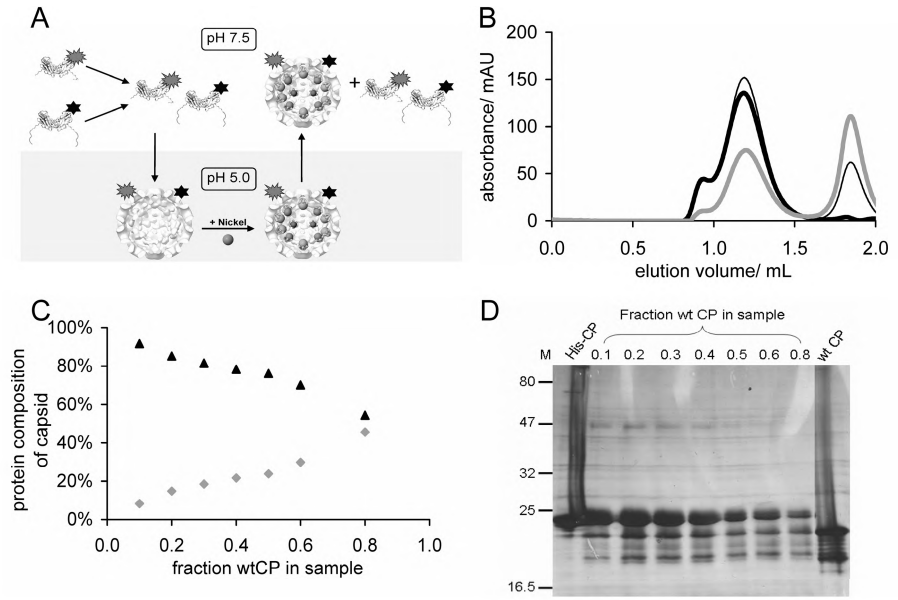

Figure 12 *Formation of capsids at pH 7.5 containing wt CP's and His-CP's. A) Schematic representation of the experiment. Wt CP was labeled with Dylight™ 647 and His-CP was labeled with Dylight™ 488. The capsid proteins were then mixed in different ratios and the pH was lowered to induce capsid formation, resulting in capsids consisting of wt CP and His-CP. Nickel chloride was added to the capsids and the pH was increased to pH 7.5 and the resulting mixtures were analyzed by FPLC B) FPLC trace of a mixture containing 30% wt CP. The thin black line represents the protein absorption at λ=280 nm, the thick black line represents the absorption of the labeled His-CP at λ=493 nm and the grey line represents the absorption of the labeled wt CP at λ=646 nm. The stabilized capsids elute at V=1.15 mL and the capsid protein dimers elute at V=1.8 mL. The FPLC system is equipped with a Superose 6 column. C) The percentages of wt CP (grey diamonds) and His-CP (black triangles) in the capsids eluting at V=1.15 mL were calculated using the absorptions of the Dylight™ labels measured on FPLC. The calculated percentages of wt CP and His-CP present in the formed capsids were plotted against the initial fraction of wt CP present in the samples. D) Silver stained SDS-Page gel of FPLC fractions of stabilized capsids at pH 7.5 composed of His-CP and wt CP. M indicates the marker, molecular weights are given in kDa. Pure His-CP and wt CP are loaded on either side of the gel.*

The fraction of wt CP in the assembled capsids could be calculated using the different absorbances of the labeled wt CP and His-CP. Since the labeling efficiency was not the same for the wt CP and the His-CP, a correction factor had to be introduced in order to compare the absorption values of the two dyes *(see experimental section for calculation details)*.

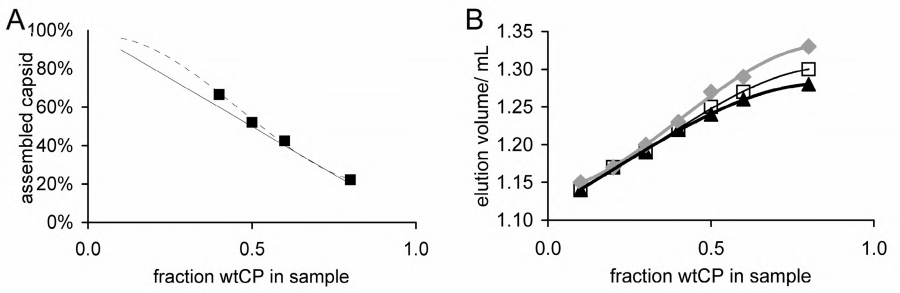

Figure 13 *Formation of capsids at pH 7.5 containing wt CPs and His-CPs. A) The amount of assembled capsid in the sample was calculated using the capsid peak and dimer peak areas of the FPLC graphs. The black squares show the percentage of capsid protein that was assembled as a function of the fraction of wt CP present in the sample. The solid line indicates the fraction of His-CP present in the samples. B) The elution volumes of the assembled capsids vary as a function of the fraction of wt CP present in the sample. Open black squares represent the elution volumes of the capsid proteins as monitored by the absorption at λ=280 nm, grey diamonds represent the elution volumes of the His-CP as monitored by the absorption at λ=493 nm, and black triangles represent the elution volumes of the wt CP as monitored by the absorption at λ=646 nm.*

The fraction of wt CP in the assembled capsids increases with increasing amounts of wt CP present in the mixture (figure 12c). However, the fraction of wt CP in the capsids was smaller than the fraction wt CP in the total mixture. This would suggest that more wt CP was in its dimeric form than His-CP. From the FPLC graphs it can clearly be seen that this is the case. The dimer peaks are almost solely composed of wt CP.

With increasing wt CP:total protein ratios, the capsid peak showed a shift to higher elution volumes. At low ratios the elution volume was V=1.14 mL which corresponds to a T=3 capsid, but at higher wt CP:total protein ratios, the elution volume shifted to V=1.30 mL, which is indicative of a smaller capsid (figure 13b). TEM micrographs confirmed this decrease in size (figure 14).

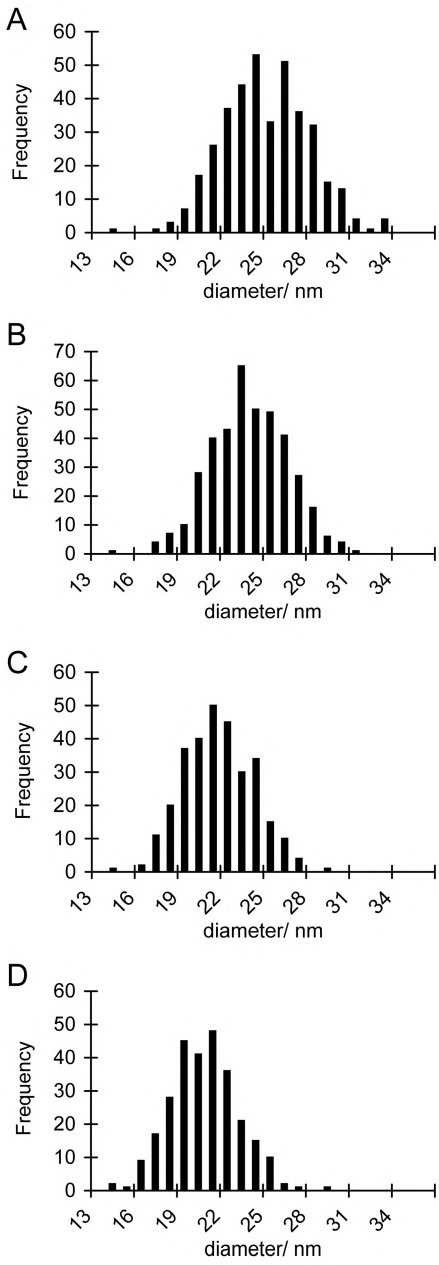

Figure 14 *Size distribution of mixed His-CP wt CP capsids stabilized with $NiCl_2$ at pH 7.5, based on TEM micrographs. A) Fraction of wt CP was 0.2. B) Fraction of wt CP was 0.4. C) Fraction of wt CP was 0.6. D) Fraction of wt CP was 0.8.*

Encapsulation of EGFP using coiled-coils in nickel stabilized capsids at pH 7.5

In a first step towards a multi-enzyme virus-based nanoreactor[12] we were able to encapsulate multiple EGFP proteins into the capsid by attaching EGFP to the capsid protein prior to pH induced assembly as described in chapter 6. The capsid protein and EGFP were genetically modified to contain a coiled-coil amino acid sequence. The coiled-coils bind non-covalently to each other, and thus the capsid protein could be attached to EGFP by mixing the proteins. Since this complex was too bulky to form capsids, the complex was mixed with non-functionalized wt CP before induction of assembly by decreasing the pH. Multiple EGFP proteins could be encapsulated this way. One major drawback of a nanoreactor formed is this way, however, is the low pH which is required to form and stabilize the capsids. Stabilization at pH 7.5 might be accomplished by combining the method to stabilize capsids using nickel, with the encapsulation strategy described in chapter 6.[12]

As a small amount of wt CP does not seem to destabilize the nickel-stabilized capsid, a small amount of EGFP-capsid protein complex might also be tolerated. So for this experiment, wt CP was replaced by the EGFP-capsid protein complex. 10% of EGFP-capsid protein complex was mixed with 90% His-CP at pH 7.5, and the assembly was induced by lowering the pH to 5.0 (figure 15a). The capsid was then separated from non-encapsulated EGFP and dimers by FPLC. Using the absorption of EGFP at λ=395 nm and the protein absorption at λ=280 nm, it could be determined that six EGFP proteins had been encapsulated per capsid *(see experimental section for calculation details)*. This procedure was repeated and the purified fractions containing the highest concentration of capsid were combined. Nickel was then added to these purified capsids and the pH was increased to 7.5. The procedure to encapsulate EGFP at pH 7.5 was also repeated without purification of the capsids containing EGFP at pH 5.0 before addition of nickel. This resulted in samples containing a larger amount of large aggregates (figure 15d) than when the capsids were pre-purified at pH 5.0.

Analysis by FPLC showed that the capsids were still largely intact and still contained 6 EGFP proteins per capsid (figure 15c and d). The capsid fraction at V=1.15 mL was analyzed by SDS-PAGE gel, and the proteins were detected by silver staining (figure 15e). This confirmed the presence of EGFP in the capsid fraction of the FPLC. Capsids containing EGFP at pH 5.0 and pH 7.5 were also analyzed by fluorescence spectroscopy. Since EGFP is pH sensitive, its quantum yield is very low at pH 5.0 and no fluorescence signal could be detected. The

capsids containing EGFP at pH 7.5, however, did give a clear fluorescence signal, clearly pointing to the presence of intact EGFP at pH 7.5 in the samples (figure 15f).

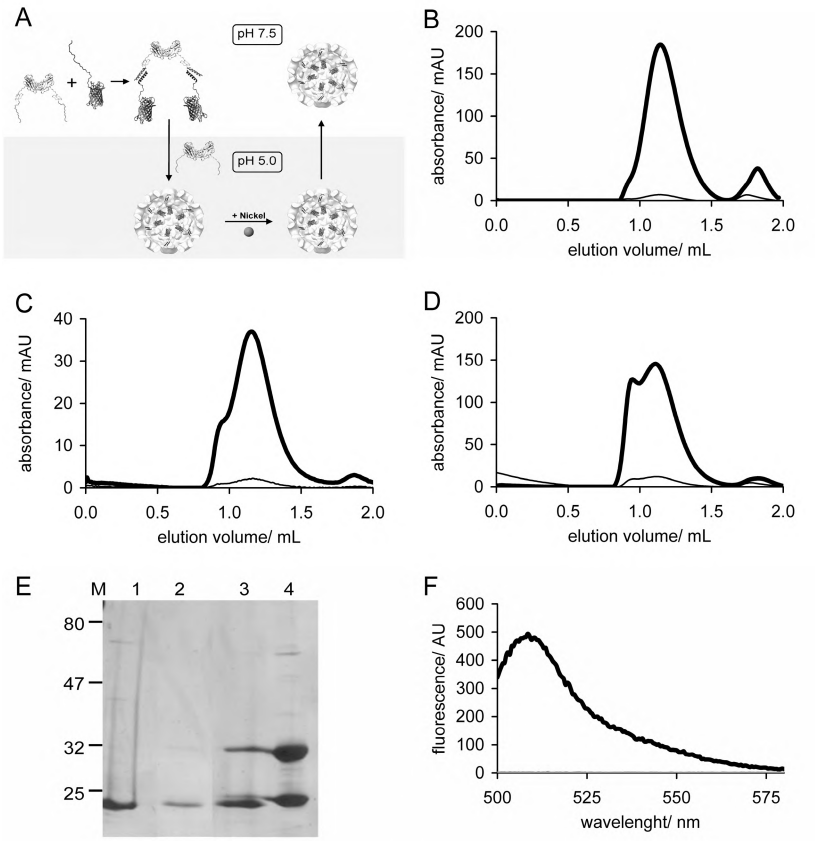

Figure 15 *Encapsulation of EGFP using coiled-coils in nickel stabilized capsids at pH 7.5. A) Schematic representation of the experiment. Capsid protein with K-coil is added to EGFP with E-coil to form the EGFP-capsid protein complex. This complex is then mixed with His-CP at pH 7.5 and subsequently dialyzed to pH 5.0 to induce capsid formation. The capsids are then either first purified on FPLC before addition of $NiCl_2$ or directly stabilized without additional purification. The pH is subsequently increased to 7.5 and the resulting mixture is analyzed by FPLC. B) FPLC trace of capsid with EGFP encapsulated at pH 5.0. The thick black line represents the protein absorption at λ=280 nm, the thin black line represents the EGFP specific absorption at λ=395 nm. The stabilized capsids elute at V=1.15 mL and the capsid protein dimers elute at V=1.8 mL. The FPLC system is equipped with a Superose 6 column. C) After addition of $NiCl_2$ to the FPLC purified capsids shown in graph B, the pH was increased to 7.5 and the capsids were reinjected on the FPLC column. The thick black line represents the protein absorption at λ=280 nm. The thin black line represents the EGFP*

specific absorption at λ=490 nm. D) FPLC trace of stabilized capsids with EGFP at pH 7.5, which were not purified prior to addition of $NiCl_2$. E) Silver stained SDS-Page gel of FPLC fractions of stabilized capsids containing EGFP at pH 7.5. M indicates the marker, molecular weights are given in kDa. Lane 1: His-CP. Lane 2: Capsids containing EGFP purified at pH 5.0 prior to stabilization (from graph 15C), due to the low concentration of this sample the EGFP is barely visible. Lane 3: Capsids containing EGFP not purified at pH 5.0 prior to purification (from graph 15D). Lane 4: GE-CK complex F) Fluorescence spectrum of EGFP containing capsids at pH 7.5. The sample was excited at λ=490 nm. The black line represents the capsids with EGFP at pH 7.5; the grey line represents the capsids with EGFP at pH 5.0.

Concluding remarks

The CCMV capsid protein is a versatile nano-building block, of which the self-assembly behavior can be tuned by changing the solvent parameters or by the addition of polyelectrolytes to form a range of different structures.

To further extend and employ the capsid proteins they were heterologously expressed in *E.coli* bacteria to allow for their genetic modification. The capsid protein produced by *E.coli* is identical to the wt CP, except for the N-terminal domain where a thrombin cleavage site and a His-tag were added. Upon lowering the pH to 5.0 the His-CP assembles into a 28 nm sized capsid, hence the addition of a His-tag does not seem to hamper the pH induced assembly. It does however expand the possibilities for assembly of the capsid protein. The His-tag is able to bind to a range of metal ions, and when these metal ions are added to a solution of His-CP at pH 7.5, a portion of the proteins assembles into particles of approximately 22 nm in size. Furthermore, when nickel ions are added to assembled His-CP capsids at pH 5.0, the interaction between the nickel and the His-tags seemingly stabilizes the capsid, such that it also remains stable when the pH is afterwards increased to 7.5. In contrast to what happens when nickel ions are added to the capsid proteins at pH 7.5, the nickel His-tag interaction at pH 5.0 presumably does not induce the assembly of the capsid, but only stabilizes the already formed capsid. The His-tags of the capsid proteins are closely located to each other at the centers of the hexamers and perhaps also at the centers of the pentamers. It is conceivable therefore that a nickel ion is coordinated to several his-tags, connecting the capsid proteins to each other. This might stabilize the hexamers, which in turn leads to stabilization of the entire capsid, as all hexamers are interconnected through dimer contacts of the C subunits.

Stabilization of the capsid at pH 7.5 could prove to be very useful in, for example, the construction of an enzyme containing nanoreactor. Enzymes are often studied in solution, while in living organisms, they reside in the closely confined and highly structured environment of cell organelles. Studying enzymes in the confined space of a nanometer-sized reaction vessel such as the CCMV capsid, mimicks the cellular environment and thus may provide fundamental insights into the way living cells work and process molecules.[13]

In chapter 6[12] it was shown that it is possible to encapsulate multiple proteins in the capsid interior, which is a first step in the development of such a nano-reactor. A major disadvantage of this system however is the low pH required to keep the capsid intact, since many enzymes do not function well at low pH. Therefore, the ability to stabilize the capsid at pH 7.5 is an important step in the realization of a multi-enzyme nanoreactor. Another important step is the ability to combine the system described in chapter 6 to encapsulate multiple proteins within the CCMV capsid, with the method described in this chapter to stabilize the capsids at pH 7.5, in order to encapsulate multiple enzymes in the nickel stabilized capsids at pH 7.5. Since proteins were non-covalently attached to the capsid proteins in order to be able to encapsulate them, a portion of the capsid needs to consist of this protein-capsid protein complex. To test whether the nickel stabilized His-CP capsid would still be stable if a portion of the capsid proteins would not contain His-tags, capsids were assembled that contained both wt CP and His-CP. This showed that a small portion of capsid proteins lacking a His-tag could be incorporated into the capsid, without significantly affecting the stability.

To create nickel stabilized capsids at pH 7.5 containing EGFP as a model protein, a small amount of EGFP-capsid protein complex was mixed with His-CP to create EGFP containing capsids at pH 5.0. After addition of nickel, these capsids were stable at pH 7.5, while still containing approximately the same amount of EGFP. Further tools for the creation of a multiple enzyme nanoreactor have now been developed using a genetically modified version of the capsid protein, which contains a metal binding N-terminal His-tag sequence. This sequence expands the assembly properties of the CCMV capsid protein. By the addition of nickel ions, both empty capsids and capsids which contain multiple encapsulated proteins can now be stabilized at a physiologically relevant pH. In the future, this technique could be used to study the effect of close confinement on enzymes at a physiologically relevant pH.

Experimental section

Materials

Protein concentrations were determined using a Cary 50 Conc (Varian, Middelburg) UV-VIS spectrophotometer using a quartz cuvette with a path length of 3 mm. FPLC measurements were performed using a Superose 6 PC 3.2/30 analytical column from GE lifesciences, on an Amersham Ettan LC system, fitted with a fractionating device. Buffers for FPLC were filtered with a Millipore 0.2 μM filter before use. TEM grids (Formvar-Carbon) were exposed to an electron discharge treatment using a Cressington Carbon coater and power unit. The sample was applied to the grids by adding a 5 μL drop of sample solution (~ 0.2 mg/mL) to the grid and carefully removing it after 1 minute immersion using a filter paper. The grid was allowed to dry for at least 15 min. before applying 5 μL of a 2% (w/v) uranyl acetate aqueous solution, which was removed after 15 s. The grid was again allowed to dry for at least 15 min. Samples were studied on a JEOL JEM-1010 TEM (Jeol, Japan). DLS measurements were performed on a Zetasizer Nano S (Malvern Instruments Ltd, England), samples were first purified by FPLC and subsequently measured. Fluorescence measurements were performed on a Perkin Elmer LS-55 fluorescence spectrometer and excited at 490 nm (slit width: 10 nm) and fluorescence emission was detected (slit width: 9 nm) from 500 to 650 nm at 10 nm/s. After purification of the proteins two buffers were used; a buffer pH 7.5 (50 mM phosphate, 0.5 M NaCl) and buffer pH 5.0 (50 mM acetate, 0.5 M NaCl, 10 mM CaCl2), the pH of the buffers was set with HCl.

Creation of a heterologous expression system for the capsid protein in E. coli cells

This experiment was described in chapter 6.

Expression of His-CP

One colony of BL21(DE3)pLysS cells expressing His-CP was used to inoculate 100 mL of LB medium containing ampicillin (0.050 g/L) and chloramphenicol (0.025 g/L). After growth overnight at 30°C this culture was used to inoculate 900 mL of LB medium containing ampicillin (0.05 g/L) and chloramphenicol (0.025 g/L) and grown at 30°C. Protein expression was induced during logarithmic growth (OD600 =0.4-0.6) by addition of IPTG to a final concentration of 1 mmol/L.

Purification of His-CP

Cells were harvested after 7 hrs of expression by centrifugation at 4000 rpm for 15 min. at 4^0C. The supernatant was discarded and the pelleted cells were stored at -20^0C. After thawing, the pelleted cells were resuspended in approximately 10 mL of lysis buffer (50 mM NaH_2PO_4, 10 mM imidazole and 1.5 M NaCl pH 8.0). Approximately 10 mg of lysozyme (EC 3.2.1.17 from Fluka) was added and the solution was incubated at 4^0C for 30 min. The solution was sonicated for 5 times 10 s with duty cycle 40 and output control 6 (Branson

Sonifier 250, marius instruments Nieuwegein, the Netherlands). RNase (10 ug/mL) and DNase (5ug/mL) were added to the solution and the mixture was incubated at 4^0C for 15 min. The solution was centrifuged at 10,000 rpm for 20-30 min. to pellet the cellular debris. The supernatant was incubated with 1 mL of Ni-NTA agarose beads for 1 h at 4^0C. The flow through was collected and the column was washed with 40 mL of wash buffer (50 mM NaH_2PO_4, 25 mM imidazole and 1.5 M NaCl pH 8.0)). The capsid protein was then eluted from the column using approximately 10 mL of elution buffer (50 mM NaH_2PO_4, 250 mM imidazole and 1.5 M NaCl pH 8.0) and the protein was collected in small fractions. The fractions were dialyzed overnight to buffer pH 7.5, to remove the excess of imidazole. The protein purity was checked with SDS-Page gel electrophoresis.

Induction of capsid like structures at pH 7.5 by nickel ions

Ten equivalents of metal ion (i.e. 10 mol metal ion per one mol His-CP) were added to a solution of His-CP in phosphate buffer pH 7.5. The metal ion was dissolved in 5% of the protein solution volume of MQ. The mixture was stirred for 1 h at RT.

Stabilization of T=3 capsid structure at pH 7.5 by nickel ions

Ten equivalents of $NiCl_2$ were added to a solution of His-CP in acetate buffer pH 5.0. The metal ion was dissolved in 5% of the protein solution volume of MQ. The mixture was stirred for 1 h at RT and the sample was subsequently dialyzed to a phosphate buffer pH 7.5 for at least 3 hrs.

ICP-MS and ICP-OES

Nickel stabilized His-CP capsids were purified using FPLC to remove non-bound nickel and non-assembled proteins. FPLC fractions eluting at V=1.2 mL were dissolved in 1% (v/v) nitric acid in water. The concentrations of Ni and S were determined on a Thermo Fisher scientific IRIS Intrepid II XDL ICP-OES machine and the concentration of Ni was verified on a Thermo Fisher scientific X-series ICP-MS machine.

Protein labeling

A solution of capsid protein (500 µL, ~ 3 mg/mL) in 0.1 M phosphate buffer pH 7.5, 0.15 M NaCl, was added to a vial containing 50 µg DyLightTM NHS-ester (488 or 549 used for His-CP and 647 for wt CP) and stirred for 1 h at RT. The non-coupled dye was removed by dialysis to a buffer pH 7.5 or pH 5.0.

Mixed His-CP wt CP capsids

Dylight labeled His-CP and wt CP in phosphate buffer pH 7.5 were mixed in the desired ratio for 10 min. at 4^0C, prior to dialysis to acetate buffer pH 5.0, to induce assembly.

Encapsulation of EGFP using coiled-coils in nickel stabilized capsids at pH 7.5

Expression and purification of the EGFP-capsid protein complex was carried out according to a literature procedure.[12]

90% His-CP was mixed with 10% EGFP-capsid protein complex for 10 min. and capsid assembly was induced by dialysis to acetate buffer pH 5.0. The capsids were then either purified on FPLC equipped with a superose 6 column to remove all non-encapsulated proteins, or directly stabilized without purification. Capsids were stabilized by addition of 10 equiv. of $NiCl_2$. This mixture was stirred for 1 h at RT and the samples were subsequently dialyzed to a phosphate buffer pH 7.5 for at least 3 hrs.

Calculations

To determine the ratio of wt CP and His-CP in the mixed nickel stabilized capsids (figure 12c), the different dye absorptions were used. Wt CP and His-CP were separately labeled with two different dyes. Wt CP was labeled with Dylight™ 647 and His-CP was labeled with Dylight™ 488. Since the proteins were not labeled with equal efficiency and the extinction coefficients of the dyes are different, the measured absorptions can not be compared directly. It is however possible to calculate a correction factor and using this correction factor the two absorptions can be compared directly, i.e. equal absorption then indicates the presence of equal amounts of protein.

Since the ratio between the His-CP and wt CP in the sample is known, the absorbance of all eluted peaks in the FPLC trace should reflect this ratio after using the correction factor. So if 90% of the sample consists of His-CP, then the total absorption of all peaks at λ=488 nm, should be 90% of the combined absorptions at λ=650 nm and λ=488 nm. The corrected absorption for wt CP would then be nine times less then this absorption, since the wt CP only constitutes 10% of the sample.

Error! Bookmark not defined. $A_{wtCP\_corrected} = \frac{A_{His-CP}}{fraction_{His-CP}} fraction_{wtCP}$

If we divide the real wt CP absorption by the correction factor, we should obtain the corrected absorption value.

$$A_{wtCP\_corrected} = \frac{A_{wtCP}}{correction\_factor}$$

So if we combine both formulas, we obtain the formula for the correction factor.

$$correction\_factor = \frac{A_{wtCP}}{\frac{A_{His-CP}}{fraction_{His-CP}} fraction_{wtCP}}$$

This correction factor can now be calculated for every sample, and the corrected absorption values can directly be used to calculate the percentage of wt CP or His-CP in the capsids, since the ratio between the absorptions of both dyes now directly reflects the ratios of the proteins.

References

1. de la Escosura, A., R.J.M. Nolte, and J. Cornelissen, *Viruses and protein cages as nanocontainers and nanoreactors.* J. Mater. Chem., **2009**, *19*, 2274.
2. Takafumi, U., et al., *Size-Selective Olefin Hydrogenation by a Pd Nanocluster Provided in an Apo-Ferritin Cage.* Angew. Chem. Int. Edit., **2004**, *43*, 2527.
3. Comellas-Aragones, M., et al., *A virus-based single-enzyme nanoreactor.* Nat. Nanotechnol., **2007**, *2*, 635.
4. Gillitzer, E., et al., *Controlled ligand display on a symmetrical protein-cage architecture through mixed assembly.* Small, **2006**, *2*, 962.
5. Carette, N., et al., *A virus-based biocatalyst.* Nat. Nanotechnol., **2007**, *2*, 226.
6. Patrick , G.H. and B.F. Matthew *Integration of a Self-Assembling Protein Scaffold with Water-Soluble Single-Walled Carbon Nanotubes.* Angew. Chem. Int. Edit., **2007**, *46*, 4370.
7. Lee, Y.J., et al., *Fabricating Genetically Engineered High-Power Lithium-Ion Batteries Using Multiple Virus Genes.* Science, **2009**, *324*, 1051.
8. Smith, T.J., et al., *The Structure of Cucumber Mosaic Virus and Comparison to Cowpea Chlorotic Mottle Virus.* Journal of Virology, **2000**, *74*, 7578.
9. Valenti, L.E., C.P. De Pauli, and C.E. Giacomelli, *The binding of Ni(II) ions to hexahistidine as a model system of the interaction between nickel and His-tagged proteins.* J. Inorg. Biochem., **2006**, *100*, 192.
10. Basu, G., et al., *Metal binding to cowpea chlorotic mottle virus using terbium(III) fluorescence.* J. Biol. Inorg. Chem., **2003**, *8*, 721.
11. Konecny, R., et al., *Electrostatic properties of cowpea chlorotic mottle virus and cucumber mosaic virus capsids.* Biopolymers, **2006**, *82*, 106.
12. Minten, I.J., et al., *Controlled Encapsulation of Multiple Proteins in Virus Capsids.* J. Am. Chem. Soc., **2009**, *131*, 17771.
13. Tanaka, S., M.R. Sawaya, and T.O. Yeates, *Structure and Mechanisms of a Protein-Based Organelle in Escherichia coli.* Science, **2010**, *327*, 81.

Chapter 9

*Catalytic capsids: the effect of confinement**

Introduction

To better understand the complex workings of the cell, it is important to appreciate the interplay between the enzyme and its environment. In the cell, enzymes are invariably spatially confined in crowded and tightly controlled cellular organelles. Even in bacteria, which do not have organelles, enzyme loaded MCPs have been recently discovered, which have a similar function. It is therefore important to study enzymes in an environment which mimics these confined conditions where normal solution reaction kinetics is no longer valid. It is, however, challenging to find suitable biomimetic containers and to entrap enzymes within these artificial systems in a controlled way. Polymeric vesicles,[1,2] mimicking lipid vesicles, are a commonly employed system for the encapsulation of enzymes as these are robust and easy to use. Despite their potential, their relatively large size (diameter (D) = ~ 100 nm) combined with usually a high polydispersity results in non-uniform reaction conditions in each vesicle hindering kinetic analysis.

In contrast, protein cages, like virus capsids, can have diameters of only tens of nanometers and are exceptionally monodisperse. Viruses are the most abundant biological entities on earth,[3] and are mostly known for their infectious properties. They often consist of only one or a few different protein components and nucleic acids, presented by Nature in numerous variations. Their simple composition makes them very useful as versatile nanobuilding blocks, e. g. for the creation of conductive nanowires,[4] MRI contrast agents[5] and drug delivery systems.[6]

One example of a frequently used virus is CCMV. Its unique assembly properties have been used to encapsulate inorganic materials,[5,7] negatively charged

* This work was published: *Chem. Sci.*, **2010**, DOI: 10.1039/C0SC00407C

polymers,[8,9] or even enzymes[10] by adding these molecules during the assembly process.

The encapsulation of single enzyme molecules inside the CCMV capsid by a statistical procedure has already been demonstrated.[10] Following up on this work, we have described a method using coiled-coils to encapsulate multiple proteins in the confined space of the CCMV capsid (see chapter 6).[11] This method significantly increases the encapsulation efficiency and, most importantly, provides control over the average number of encapsulated proteins. This directed assembly process closely matches the strategy used by Nature to pack enzymes into MCPs via short N-terminal peptide sequences.[12] In this chapter, we aim to use this method to encapsulate the enzyme Pseudozyma (formerly Candida) antarctica lipase B (PalB) and to investigate the overall reaction rate as a function of the number of encapsulated PalB molecules. In addition we co-encapsulated EGFP as a non-catalytic protein to investigate the reaction rate as a function of the total protein concentration inside the capsid.

Results and Discussion

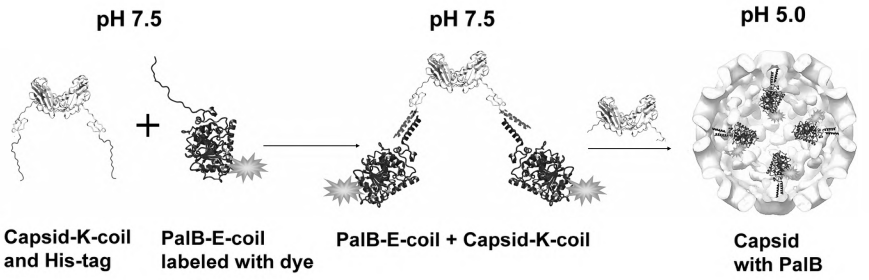

Figure 1 *Schematic representation of PalB encapsulation. Capsid with K-coil and His-tag will be referred to as capsid protein in the text, PalB with E-coil will be referred to as PalB in the text. The PalB-capsid protein complex is mixed with wild-type capsid at pH 7.5 and subsequently dialyzed to pH 5.0 to induce capsid formation.*

PalB (EC 3.1.1.3), a widely used biocatalyst, was chosen as the model enzyme since it retains most of its activity at pH 5.0.[13] PalB was genetically modified with the E-coil monomer of the heterodimeric coiled coil, while the capsid protein was modified with the complementary K-coil (see chapter 6). To enable the encapsulation efficiency to be determined, PalB was labeled with the fluorescent dye Alexa Fluor 568. Coiled-coil formation was achieved by simply mixing capsid

protein with PalB. The resulting complex was purified by size exclusion chromatography and subsequently mixed with wildtype capsid protein in various ratios to form capsids with different amounts of encapsulated PalB. This capsid assembly was induced by lowering the pH to 5.0 (figure 1). Following assembly, the capsids were purified and analyzed by size exclusion chromatograpy using a Superose 6 column (figure 2) and TEM (figure 3). The UV-Vis absorption of the protein and the dye was monitored during the run and the ratio between the dye-specific and protein-specific absorption of the non-encapsulated CalB was used to determine the amount of CalB encapsulated per capsid (see experimental section). Four different batches of capsid were made, with on average 1.3, 2.0, 3.5 and 4.0 CalB enzymes per capsid.

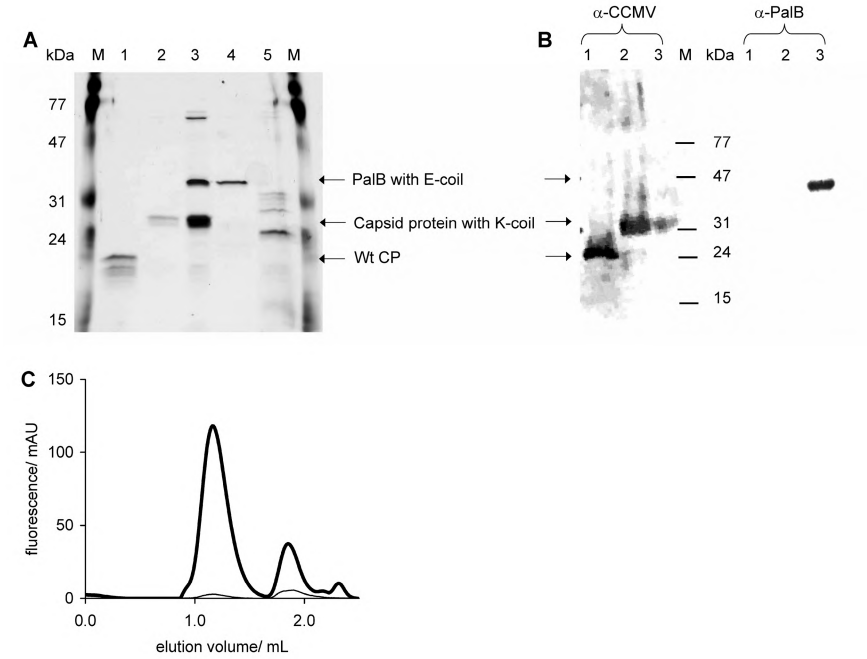

Figure 2 *Analysis of proteins and capsids. A) Silver stained SDS-PAGE gel of proteins and protein complexes used in the experiments. M indicates the marker, molecular weights are given in kDa. Lane 1: wt CP. Lane 2: Capsid protein with K-coil. Lane 3: PalB-capsid protein complex after purification on FPLC. Lane 4: PalB with E-coil. Lane 5: GE-CK complex. B) Immuno blots. Lane 1: wt CP. Lane 2: Capsid protein with K-coil. Lane 3: Capsid protein with K-coil and PalB with E-coil. Left blot was probed with anti-CCMV antibody, right blot was probed with anti-PalB antibody.*

C) FPLC trace of a mixture of PalB-capsid protein complex with wt CP at pH 5.0. The FPLC system is equipped with a Superose 6 column. The thick black line represents the protein absorption at λ=280 nm, and the thin black line the absorption of the dye with which PalB is labeled at λ=586 nm. The capsid with encapsulated PalB elutes at V=1.1 mL, and the unassembled capsid dimers and of PalB-capsid protein complex elute at V ~1.8 mL.

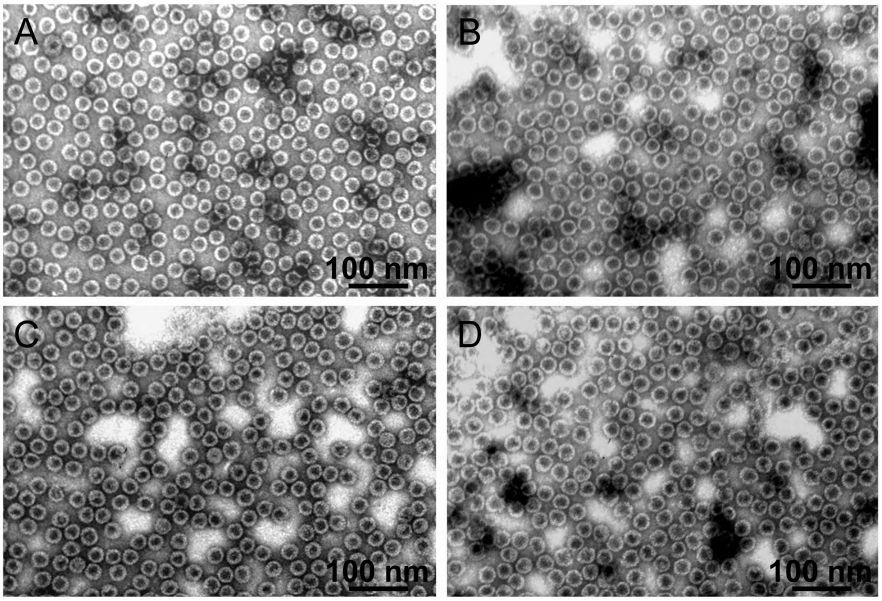

Figure 3 *TEM micrographs of uranyl acetate stained samples. A) Empty wild-type capsids. B) Capsids filled with 1 PalB per capsid C) Capsids filled with 4 PalB per capsid. D) Capsids filled with 1 PalB and 7 EGFP per capsid.*

In steady-state enzyme kinetic measurements, when the concentration of the substrate is well above K_M, the enzyme functions at its maximum rate (k_{cat}). The reaction rate is then directly proportional to the enzyme concentration in solution. To test if this is also true for our confined system, the four samples containing different amounts of encapsulated PalB and a sample with non-encapsulated PalB were diluted such that the concentration of PalB was the same for all samples (figure 5a). To determine the overall reaction rate of these different samples, the pro-fluorescent substrate 6,8-difluoro-4-methylumbelliferyl octanoate (DiFMU-octanoate) was added to give a final concentration of 40 μM, which is well above the K_M values of 8.8 μM determined in this study (see figure 4a and experimental

section) and 2.6 μM reported for PalB and DiFMU octanoate in the literature.[14,15] The initial reaction rates were calculated and normalized with respect to the sample with the lowest encapsulated enzyme concentration (see figure 5b and experimental section). Since the overall enzyme concentration is the same in all five samples, the reaction rates should also be the same, if these were to follow standard Michaelis-Menten reaction kinetics. We observed, however, that the reaction rate is inversely proportional to the number of encapsulated enzymes (figure 5c) and the encapsulated enzymes appeared to outperform the non-encapsulated PalB.

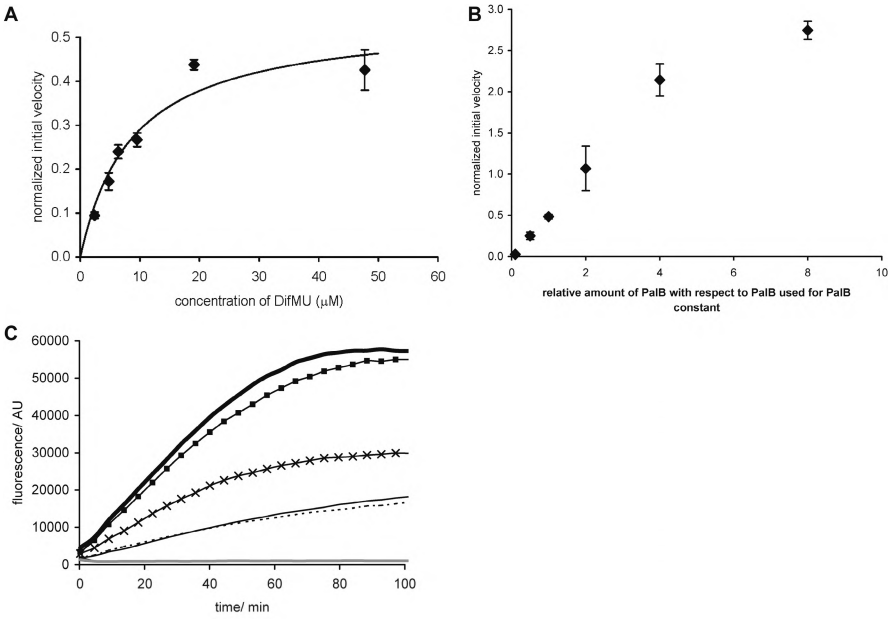

Figure 4 *A) Michaelis-Menten saturation curve of the normalized initial reaction rate versus the DiFMU concentration. The K_M value calculated from this data is 8.8 μM. B) Plot of the normalized initial rate versus PalB concentration. DiMFU concentration was 40 μM. C) DiFMU substrate conversion curves. The concentration of PalB was the same in all samples. Grey line depicts buffer with DiFMU, thin dotted line depicts the control sample with non-encapsulated PalB, solid thin line depicts sample with 4 PalB/capsid, thin line with crosses depicts sample with 3.5 PalB/capsid,thin line with squares line depicts sample with 2 PalB/capsid, and thick black line depicts sample with 1.3 PalB/capsid.*

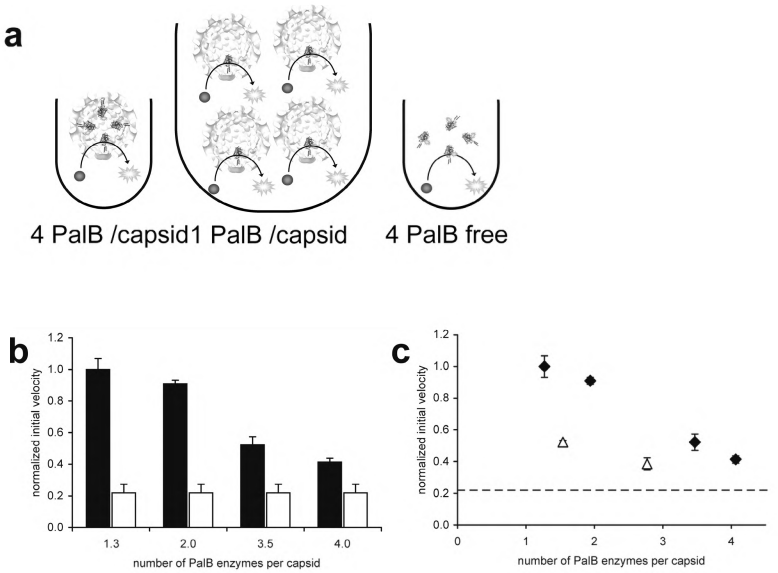

Figure 5 *Kinetics of encapsulated PalB. The concentration of PalB is the same in all samples. A) Schematic representation of experimental setup. B) Initial rates normalized with respect to 1.3 PalB/capsid. Black bars represent initial rates of samples with encapsulated PalB. White bars indicate normalized initial velocity of non-encapsulated PalB (one experiment was plotted four times for clarity.) Error bars indicate SD of triplicate experiments. C) Normalized initial rates plotted against the number of encapsulated PalB enzymes per capsid. Filled diamonts represent capsid with encapsulated PalB; open triangles represent capsids with the indicated number of PalB and 7 EGFP proteins. Error bars indicate SD of triplicate experiments. The dashed line indicates the initial velocity of non-encapsulated PalB.*

It should be noted that the overall amount of capsids is higher in the samples with less PalB per capsid. In order to eliminate any effect of the capsid on the catalytic reaction, four samples containing different amounts of PalB were diluted such that all samples contained the same amount of capsids. Since the concentration of PalB is now different for each sample, four different control experiments were performed, in which the concentration of non-encapsulated PalB was the same as for the corresponding encapsulated PalB samples (figure 6a). As expected, the initial reaction rates of non-encapsulated PalB increase with increasing enzyme concentrations (figure 6b). The initial reaction rates of the encapsulated PalB samples, however, are relatively constant within experimental error.

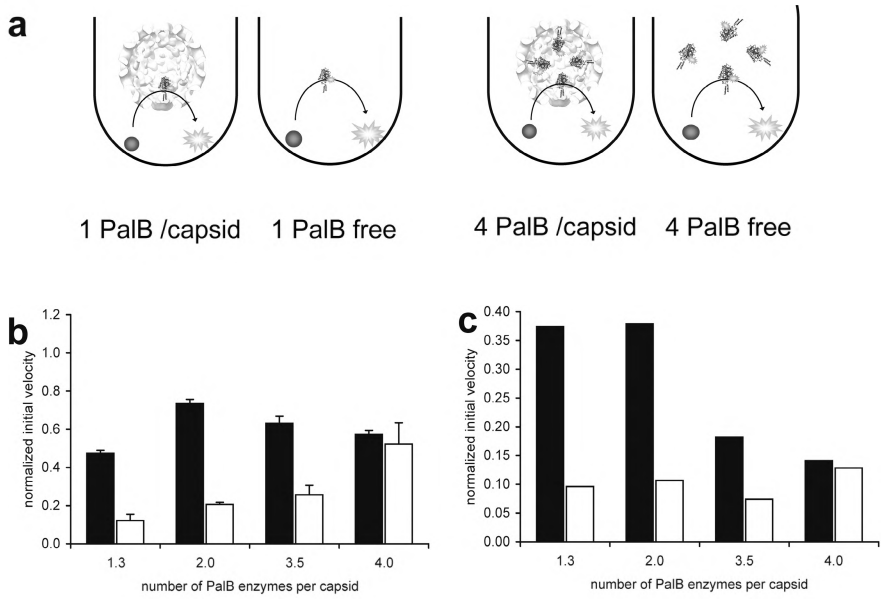

Figure 6 *Kinetics of encapsulated PalB. The concentration of capsid is the same in all samples. A) Schematic representation of the experimental setup. B) Initial rates normalized with respect to 1.3 PalB/capsid experiment reported in figure 5b. The reduced normalized rate with respect to figure 5b is a result of the absolute dilution in order to obtain the same concentration of capsid in all samples. Black bars represent initial rates of samples with encapsulated PalB. White bars indicate the normalized initial rate of non-encapsulated PalB. Error bars indicate standard deviation of triplicate experiments. C) Normalized initial rates per PalB and per capsid; values are obtained by dividing the values of the initial rates of graph B by the number of PalB enzymes per capsid.*

Dividing the activity of the non-encapsulated PalB by its concentration (figure 4c), the calculated activity should be the same for the four different samples. Within experimental margins this is the case. The calculated activity of the encapsulated PalB, however, appears to be the largest when only one or two PalB molecules are encapsulated per capsid, suggesting that an increase in the protein concentration inside the capsid might have a negative effect on the reaction rates due to some form of crowding effect.

In summary, these results show two remarkable effects of the encapsulation of PalB in the very confined space of the CCMV capsid, i.e. (i) encapsulated PalB seems to outperform the non-encapsulated PalB, and (ii) the presence of more PalBs per capsid does not lead to an increase in the overall reaction rate. The former of the two effects could be the result of a change in k_{cat} upon encapsulation inside the virus. It is known that PalB is more active when immobilized, and one could argue that the enzyme is immobilized within the capsids. Furthermore, Chen et al.[16] have shown that in vesicle systems the turnover number (k_{cat}) of entrapped enzymes increases when the vesicle size becomes smaller. They observed a 2-fold increase in turnover number going from non-encapsulated enzymes to enzymes spatially confined within vesicles with an internal volume of 5 atto liter (5×10^{-18} L), and an additional 2-fold increase going from these vesicles to vesicles with an internal volume of 0.05 atto liter. In contrast, the second effect (ii) could be explained by a crowding effect inside the virus. This potential contribution of crowding will be discussed in more detail below.

Both effects together can be readily explained by a change in the effective molarity of the enzyme in the confined space of the capsid, which we define as confinement molarity M_{conf}. The internal volume of the CCMV capsid is approximately 5 zepto liter (5×10^{-21} L), a factor 10 smaller than the smallest vesicles tested in the paper by Chen et al. In this small volume M_{conf} for the encapsulated enzyme is ~1mM, which is a factor 10^4 higher than the bulk enzyme concentration (30-100 nM). Assuming that the rate of diffusion through the capsid pores is significantly larger than the rate of catalysis and that the concentration of substrate inside the capsid and outside the capsid is equal (steady state situation) we can calculate that the enzyme concentration inside the capsid is approximately 25 times higher than the substrate concentration of 40 μM. Considering that K_M is not altered as a result of the encapsulation, the high M_{conf} can therefore be expected to shift the equilibrium towards the formation of the enzyme-substrate complex thereby increasing the reaction rate. The observed results clearly highlight that the advantage of confinement is the shifting of the enzyme-substrate equilibrium completely in the direction of the enzyme-substrate complex. Faster formation of the enzyme substrate complex might be further assisted by an increased number of collisions between the enzyme and the substrate resulting from the confinement as has been shown by *Chen et al.* [16]

Having explained the overall increase in reaction rate, we can now look at the number of molecules inside the capsid in more detail to explain why the reaction

rate does not increase with an increasing number of enzymes in the capsid. Using a Poisson distribution the chance that two or more substrates are present in the capsid at the same time can be calculated. This chance is 0.64%, which is negligible low. In fact, only ~10% of the capsids do contain a substrate molecule at all. It therefore makes no difference whether a capsid contains one or more enzymes, as there is maximally only one substrate molecule present. In other words, a consequence of the high M_{conf} is that upon going from 1 to 4 enzymes in the capsid the change in M_{conf} is so small that it has a minimal effect on the overall reaction rate. It should be noted that in solution the substrate concentration is in excess whereas in the capsid the enzyme concentration is in excess.

A possible effect of crowding within the capsid upon increasing in the number of enzymes cannot be fully discarded. To investigate a potential crowding effect, EGFP was co-encapsulated with PalB, which was labeled with a fluorescent dye. This protein does not contribute to product conversion. Two samples were prepared, one containing 1 PalB and 7 EGFP per capsid and the other containing 3 PalB and 7 EGFP per capsid. Co-encapsulation of the enzyme and the protein was demonstrated by Dual-Color Fluorescence Cross-Correlation Spectroscopy (FCCS). Control experiments of capsid containing exclusively PalB or EGFP showed only auto-correlation in the appropriate channel and little cross-correlation, while analysis of the samples containing both proteins clearly showed cross-correlation (figure 7) proving that indeed both biomacromolecules are present in the same capsid. From the FCCS data a typical dwell time of 2 ms was calculated. Assuming a 250 nm wide focal volume of the microscope, a diffusion constant of 7.4 $\mu m^2/s$ can be calculated for the CCMV capsid. This corresponds with a Stokes-Einstein radius of 29 nm, which is in reasonable agreement with the capsid size known from TEM images.

Co-encapsulation of EGFP did indeed result in a small decrease in the reaction rate (figure 5c), but since the total protein concentration was approximately two times higher than the highest amount of PalB encapsulated, the effect should have been more pronounced if crowding due to the high local enzyme concentration would really have been the only cause for the decrease in activity. Instead the observed lower reaction rate in figure 5c might be attributed to a decreased diffusion speed of the substrate and product molecules over the capsid shell. The EGFP and PalB molecules are attached to the N-termini of the capsid proteins and thus located in the vicinity of the pores in the capsid shell.

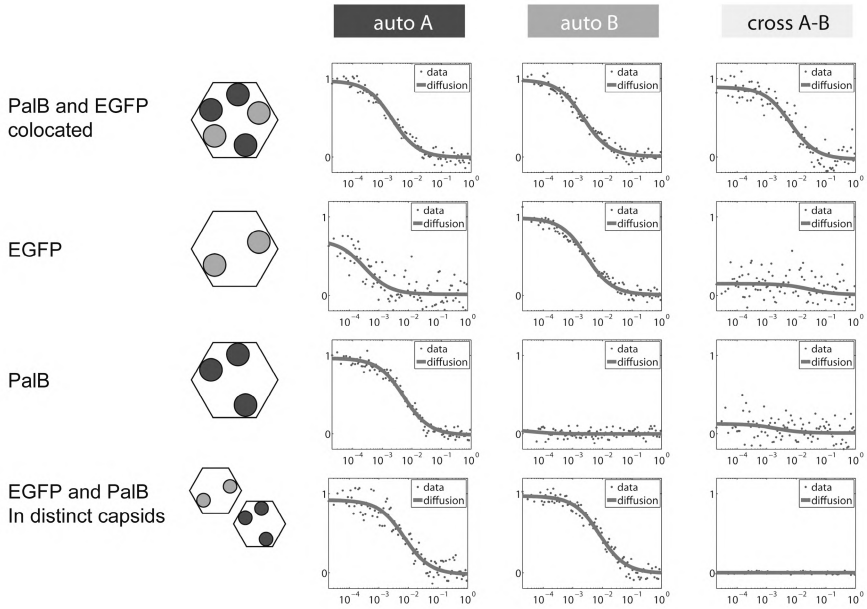

Figure 7 *Dual color FCCS correlation curves of CCMV capsids containing EGFP (grey circles), PalB (black circles) or both. When EGFP is encapsulated, there is correlation in the green channel B, while when PalB is encapsulated there is correlation in the red channel A. but only when both compounds are present in the same capsid the cross-correlation arises. The EGFP fluorescence extends to the red wavelengths, so some of the emission of EGFP is also recorded in the red channel, thereby causing some red channel auto-correlation even when the red dye Alexa Fluor 568 is absent.*

There are 20 pores of D = 2 nm in the capsid shell, together covering roughly 1% of the solvent-exposed outer surface of the capsid. Therefore diffusion over the capsid wall is likely slower than diffusion in solution.[8] If diffusion over the capsid wall would be slower than the substrate conversion, diffusion could be the rate-limiting step. This is unlikely however, since substrate conversion of native PalB is about nine orders of magnitudes slower than unhindered diffusion of a substrate molecule over the capsid wall. This diffusion rate however could be affected by an increasing number of proteins close to the pores, which will then result in a reduced overall reaction rate, as is observed (figure 5c).

Conclusion

We have studied the effects of encapsulation of multiple PalB enzymes in the spatial confinement of the CCMV capsid. The spatially confined PalB seems to have a higher activity than non-encapsulated PalB (figure 8). Although this might be caused by the immobilization of PalB in the capsid viz. a chance in k_{cat}, there is no evidence for this. It is more likely that the observed effect is caused by the high effective concentration of the enzyme combined with an increased chance of collisions due to the spatial confinement. Since a capsid rarely contains more than one substrate molecule, one enzyme is sufficient to convert the substrate, and the presence of additional enzymes has only a minimal effect on the reaction rate. Further research will be needed to confirm this theory by varying the substrate concentration, which in the described system however is hindered by the limited solubility of the DiFMU.

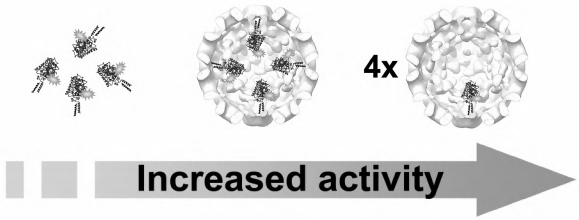

Figure 8 *Scheme representing the main findings of this chapter.*

We have shown in this chapter that it is indeed possible to encapsulate two different proteins inside the CCMV capsid, i.e. EGFP and PalB. In the future multiple different enzymes, capable of performing cascade reactions might be incorporated in the CCMV capsid using this method. When multiple different enzymes are encapsulated in the CCMV capsid it can readily be predicted that due to M_{conf} the reaction rate of the cascade reactions will be considerably enhanced. The converted substrate of one encapsulated enzyme would more likely collide and bind with another enzyme in the capsid than diffuse through the capsid wall. Thus the converted substrate would react with a different enzyme present in the same capsid at a much faster rate than if the enzymes were not encapsulated. The study of cascade reactions in confinement could not only potentially provide insight into the mechanisms of enzymatic cascade reactions inside cellular organelles, but also lead to the creation of a highly efficient nano-sized reaction vessels with potential industrial applications.

Experimental section

Materials

Protein concentrations were determined using a Cary 50 Conc (Varian, Middelburg) UV-VIS spectrophotometer using a quartz cuvet with a path length of 3 mm.

FPLC measurements were performed using either a Superose 6 PC 3.2/30 analytical column or a Superdex 200 PC 3.2/30 analytical column from GE lifesciences, on an Amersham Ettan LC system, fitted with a fractionating device. Buffers for FPLC were filtered with a Millipore 0.2 μM filter before use. TEM grids (Formvar-Carbon) were exposed to an electron discharge treatment using a Cressington Carbon coater and power unit. The sample was applied to the grids by adding a 5 μL drop of sample solution (~ 0.2 mg/mL) to the grid and carefully removing it after 1 minute immersion using a filter paper. The grid was allowed to dry for at least 15 min. before applying 5 μL of a 2% (w/v) uranyl acetate aqueous solution, which was removed after 15 s. The grid was again allowed to dry for at least 15 min. Samples were studied on a JEOL JEM-1010 TEM (Jeol, Japan).

DLS measurements were performed on a Zetasizer Nano S (Malvern Instruments Ltd, England),

Restriction enzymes and Antartic Phosphatase were obtained from New England Biolabs. DNA containing the coiled-coil sequences was obtained from Biolegio (Nijmegen, The Netherlands).

Fluorescence for enzyme kinetics was measured in 96-well flat-bottom microplates on a Wallac Victor 1420 multilabel counter.

Two buffers were used, a buffer pH 7.5 (0.5 M NaCl, 0.05 M Tris-HCl, 0.01 M $MgCl_2$ and 0.001 M EDTA) and buffer pH 5.0 (0.5 M NaCl, 0.05 M $NaCH_3COO$, 0.01 M $MgCl_2$ and 0.001 M EDTA), the pH of the buffers was set with HCl.

Capsid protein with K-coil and His-tag construction in vector pET-15b

The insert containing the K-coil DNA sequence was designed to have a 5' NdeI and a 3' NdeI restriction site overhang after annealing of the strands (table 1). The pET-15b vector containing the capsid protein sequence was digested sequentially with NdeI, dephosphorylated with Antartic Phosphatase, and purified by agarose gel elecrophoresis. The vector was mixed with the annealed insert and ligated with T4 DNA ligase. The resulting plasmids were redigested with NdeI. The insert was designed such that after ligation the plasmid would not contain NdeI restriction sites anymore. Thus ensuring that only plasmids with the insert would not be liniarized by the digestion. The plasmids were then transformed into *E. coli* XL1-Blue cells, the DNA was extracted and the sequence of the capsid protein with K-coil and his-tag was confirmed by DNA sequencing. The plasmids were then transformed into *E. coli* BL21(DE3)pLysS cells, which were used for production of the capsid protein with the K-coil and His-tag.

Table 1 *DNA and amino acid sequences. DNA sequences for the inserts used for the creation of the His-CK and PalBE proteins. Overhangs are complementary with the overhangs in the vector sequences, created with restriction enzymes. Amino acid sequences for wt CP, His-CK and PalBE protein. K-coil and E-coil amino acid sequences are printed in bold lettering.*

First line: K-coil forward (5'→3'), second line: K-coil reverse (3'→5')

```
TAGCAAAATTGCCGCGCTGAAAGAAAAAATTGCGGCCCTGAAAGAAAAAATCGCGGCGCTGAAAGAAGG
  CGTTTTAACGGCGCGACTTTCTTTTTTAACGCCGGGACTTTCTTTTTTAGCGCCGCGACTTTCTTCCAT
```

First line: E-coil forward (5'→3'), second line: E-coil reverse (3'→5')

```
TCGACGAAATTGCGGCGCTGGAAAAAGAAATTGCGGCGCTGGAAAAAGAAATTGCGGCGCTGGAAAAACTGGTGCCGCGCGGCAGCG
    GCTTTAACGCCGCGACCTTTTTCTTTAACGCCGCGACCTTTTTCTTTAACGCCGCGACCTTTTTGACCACGGCGCGCCGTCGCAGCT
```

Wt CP amino acid sequence

```
MSTVGTGKLTRAQRRAAARKNKRNTRVVQPVIVEPIASGQGKAIKAWTGYSVSKWTASCAAAEAKVTSAITISLPNELSSERNKQ
LKVGRVLLWLGLLPSVSGTVKSCVTETQTTAAASFQVALAVADNSKDVVAAMYPEAFKGITLEQLAADLTIYLYSSAALTEGDVI
VHLEVEHVRPTFDDSFTPVY
```

Capsid with His-tag and K-coil amino acid sequence

```
MGSSHHHHHHSSGLVPRGSHSKIAALKEKIAALKEKIAALKEGMMSTVGTGKLTRAQRRAAARKNKRNTRVVQPVIVEPIASGQG
KAIKAWTGYSVSKWTASCAAAEAKVTSAITISLPNELSSERNKQLKVGRVLLWLGLLPSVSGTVKSCVTETQTTAAASFQVALAV
ADNSKDVVAAMYPEAFKGITLEQLAADLTIYLYSSAALAEGDVIVHLEVEHVRPTFDDSFTPVY
```

PalB with E-coil amino acid sequence

```
MKYLLPTAAAGLLLLAAQPAMAMGLPSGSDPAFSQPKSVLDAGLTCQGASPSSVSKPILLVPGTGTTGPQSFDSNWIPLSAQLGY
TPCWISPPPFMLNDTQVNTEYMVNAITTLYAGSGNNKLPVLTWSQGGLVAQWGLTFFPSIRSKVDRLMAFAPDYKGTVLAGPLDA
LAVSAPSVWQQTTGSALTTALRNAGGLTQIVPTTNLYSATDEIVQPQVSNSPLDSSYLFNGKNVQAQAVCGPLFVIDHAGSLTSQ
FSYVVGRSALRSTTGQARSADYGITDCNPLPANDLTPEQKVAAAALLAPAAAAIVAGPKQNCEPDLMPYARPFAVGKRTCSGIVT
PLDEIAALEKEIAALEKEIAALEKLVPRGSVEHHHHHH
```

PalB with E-coil construction in vector pET-22b

The insert containing the E-coil DNA sequence was designed to have a 5' XhoI and a 3' XhoI restriction site overhang after annealing of the strands (table 1).

The pET-22b vector containing the PalB protein sequence was digested with XhoI, dephosphorylated with Antartic Phosphatase, and purified by agarose gel elecrophoresis. The vector was mixed with the annealed insert and ligated with T4 DNA ligase. The resulting plasmids were redigested with XhoI. The insert was designed such that after ligation the plasmid would not contain XhoI restriction sites anymore. Thus ensuring that only plasmids with the insert would not be liniarized by the digestion.

The resulting plasmids were transformed into *E. coli* XL1-Blue cells, the DNA was extracted and the sequence of the PalB with E-coil was confirmed by DNA sequencing. The plasmids were then transformed into *E. coli* BL21(DE3)pLysS cells, which were used for production of the PalB enzyme with the E-coil.

Expression of capsid protein with K-coil and His-tag

One colony of BL21(DE3)pLysS cells expressing His-CK was used to inoculate 100 mL of LB medium containing ampicilin (0.050 g/L) and chloramphenicol (0.025 g/L). After growth overnight at 30°C this culture was used to inoculate 900 mL of LB medium containing ampicilin (0.05 g/L) and chloramphenicol (0.025 g/L) and grown at 30°C. Protein expression was induced during logarithmic growth (OD_{600} =0.4-0.6) by addition of IPTG to a final concentration of 1 mmol/L. After induction the culture was grown at 30°C for 5 hrs.

Expression of PalB with E-coil

One colony of BLR (DE3) pLysS cells expressing PalB with E-coil was used to inoculate 100 mL of 2xTY medium containing ampicilin (0.05 g/L) and tetracycline (0.0125 g/L). After growth overnight at 30°C this culture was used to inoculate 100 mL of 2xTY medium containing ampicilin (0.05 g/L) and tetracycline (0.0125 g/L) and grown at 30°C. Protein expression was induced during logarithmic growth (OD_{600} =0.4-0.6) by addition of IPTG to a final concentration of 1 mmol/L. After induction the culture was grown at 25°C for 20 hrs.

Purification of capsid protein with K-coil and His-tag

E. coli cells containing the capsid protein with K-coil and His-tag were harvested by centrifugation at 4000 g for 15 min. at 4°C. The supernatant was discarded and the pelletted cells were stored at -20°C. After thawing, the pelletted cells were resuspended in approximately 10 mL lysis buffer (50 mM NaH_2PO_4, 10 mM imidazole and 2000 mM NaCl pH 8.0). Approximately 10 mg of lysosyme (EC 3.2.1.17 from Fluka) was added and the solution was incubated at 4°C for 30 min. The solution was sonicated for 5 times 10 s with duty cycle 40 and output control 6 (Branson Sonifier 250, marius instruments Nieuwegein, the Netherlands). RNase (10 µg/mL) and DNase (5 µg/mL) was added and the mixture was incubated at 4°C for 15 min. The solution was centrifuged at 10,000 rpm for 20-30 min. to pellet the cellular debris. The supernatant of was incubated with 0.5 mL of Ni-NTA agarose beads for 1 h at 4°C. The flow-through was collected and the column was washed with 20 mL of wash buffer (50 mM NaH_2PO_4, 25 mM imidazole and 2000 mM NaCl pH 8.0). The capsid protein was eluted from the column using approximately 10 mL of elution buffer (50 mM NaH_2PO_4, 250 mM imidazole and 2000 mM NaCl pH 8.0) and the protein was collected. The capsid protein was dialyzed overnight to buffer pH 7.5, to remove the excess of imidazole. The protein purity was checked with SDS-PAGE (figure 2).

Purification of PalB with E-coil

E. coli cells containing the PalB with E-coil enzymes were harvested by centrifugation at 4000 g for 15 min. at 4°C. The supernatant was discarded and the pelletted cells were stored at -20°C. After thawing, the pelletted cells were resuspended in approximately 10 mL lysisbuffer (50 mM NaH_2PO_4, 10 mM imidazole and 300 mM NaCl pH 8.). Approximately 10 mg of lysosyme (EC 3.2.1.17 from Fluka) was added and the solution was incubated at

4°C for 30 min. The solution was sonicated for 5 times 10 s with duty cycle 40 and output control 6 (Branson Sonifier 250, marius instruments Nieuwegein, the Netherlands). The solution was centrifuged at 10,000 rpm for 20-30 min. to pellet the cellular debris. The supernatant of was incubated with 0.5 mL of Ni-NTA agarose beads for 1 h at 4°C. The flow-through was collected and the column was washed with 20 mL of wash buffer (50 mM NaH_2PO_4, 20 mM imidazole and 300 mM NaCl pH 8.0). PalB was eluted from the column using approximately 10 mL of elution buffer (50 mM NaH_2PO_4, 250 mM imidazole and 300 mM NaCl pH 8.0) and the protein was collected. PalB was dialyzed overnight to phosphate buffer pH 8.5 (100 mM phosphate, 150 mM NaCl), to remove the excess of imidazole. The protein purity was checked with SDS-PAGE (figure 2).

Production and purification of wt CP

The purification of the CCMV virus and the removal of its RNA were carried out according to literature procedures[17,18].

Formation and purification of PalB-capsid protein complex

PalB with E-coil was labeled with Alexa Fluor® 568. Excess dye was removed by dialyses to buffer pH 7.5. Approximately 1/3 of the obtained labeled PalB with E-coil was further purified by FPLC equipped with a Superdex 200 column. The rest of the labeled PalB with E-coil was added to an two times excess of His-CK and stirred for at least 16 h at 4°C to form the PalB-capsid protein complex. The PalB-capsid protein complex was further purified by FPLC equipped with a Superdex 200 column to remove any non-complexed PalB, capsid protein or other proteins. Pure PalB with E-coil eluted at V=1.6 mL as detected at λ=280 and 586 nm or, whereas the PalB-capsid protein complex eluted between V=1.3-1.5 mL.

PalB with E-coil encapsulation

PalB-capsid protein complex in buffer pH 7.5 (50 mM Tris-HCl, 500 mM NaCl, 10 mM $MgCl_2$, 1 mM EDTA pH 7.5) was added in different ratios to wild type capsid protein in the same buffer. For the co-encapsulation of EGFP and PalB, GE-CK complex[11] was also added to the mixture. The proteins were allowed to mix for 5 min., before dialysis of the mixture overnight to buffer pH 5.0. The mixtures and non-encapsulated PalBE were then purified and analyzed on an FPLC system equipped with a Superose 6 column.

Sample preparation

Fusion proteins consisting of monomers of the coiled coil motif and PalB (E-coil), EGFP (E-coil) or the capsid protein (K-coil) were expressed in *E. coli* and purified by Ni^{2+}-NTA chromatography using standard protocols. The PalB fusion protein was fluorescently labeled with NHS-Alexa Fluor 568. Prior to capsid assembly, coiled coil complexes were formed between PalB and capsid protein as well as between EGFP and capsid protein. Complexes were mixed with wildtype capsid protein in the desired ratios in pH 7.5 buffer. Lowering the

pH to 5.0 by dialysis induced capsid assembly. Characterization of the coiled coil complexes as well as of the assembled capsids were performed using size exclusion chromatography. The absorption at 280 nm or 586 nm was used to determine the concentration of capsid protein or labeled PalB, respectively, neglecting the small contribution of PalB to the absorption at 280 nm.

Calculations of PalB encapsulation

The number of PalB enzymes in a capsid can be calculated similarly to the number of EGFP molecules in a capsid (chapter 6)[11], by using the absorption ratios of the protein absorption (λ=280 nm) and the dye absorption (λ=586 nm) at the capsid peak (V=1.1 mL) from the FPLC data.

Enzyme activity measurements

For the enzyme activity measurements with 6,8-difluoro-4-methylumbelliferyl octanoate (DiFMU-octanoate) either the concentration of PalB or the concentration of capsid was kept constant for all measured samples. Reactions were performed in triplicate using a microplate reader (λ_{ex} = 355 nm and λ_{em}=460 nm). The reaction was started with the addition of DiFMU-octanoate. The initial reaction rates were calculated from the obtained data by determining the slope of the initial part of the curves. These values were then normalized by dividing all initial slopes by the same value (i.e. the obtained initial velocity for 1.3 PalB/capsid were the concentration of PalB was kept constant throughout the experiment).

Dual-color fluorescence cross-correlation spectroscopy (FCCS)

This work will be described in detail elsewhere: V.I.Claessen, *Thesis*, Radboud University Nijmegen.[19] Laser light of 488 nm (Spectra-Physics 2080 argon ion laser) and 568 nm (Coherent CR-599 dye laser, pumped by 532 nm Spectra-Physics Millennia Nd:YAG laser) was coupled into a single-mode optical fiber, reflected by a dichroic beam splitter (Chroma z488/568rpc) and focused onto the sample by an oil immersion 100x objective (Carl Zeiss, NA = 1.30), which was mounted on a Carl Zeiss Axiovert 200 inverted microscope. Fluorescent light coming from the focal volume was collected by the same objective, passed through the dichroic beamsplitter, focused through a 100 µm pinhole, split by a second dichroic mirror (Chroma 560dcxr), filtered by either a 'red' (Chroma D630/60m) or 'green' filter (Chroma HQ525/50m) and focused onto avalanche photodiodes (PerkinElmer SPCM-AQR-14). The photon count signals were recorded using a Picoquant PicoHarp 300E TCSPC module.The focal volume was positioned in a solution of CCMV capsids containing EGFP, PalB or both. EGFP is autofluorescent and is excited at 488 nm, the PalB was labeled with an Alexa568 dye which is excited at 568 nm. Emission of both fluorophores was filtered and recorded by separate photodetectors. The photon correlation functions were calculated using matlab scripts developed in-house. Curves were normalized on the G(0) parameter and then multiplied by the r-square of the fit for clarity.

References

1. van Dongen, S.F.M., et al., *A Three-Enzyme Cascade Reaction through Positional Assembly of Enzymes in a Polymersome Nanoreactor.* Chem.-Eur. J., **2009**, *15*, 1107.
2. Caruso, F., et al., *Enzyme Encapsulation in Layer-by-Layer Engineered Polymer Multilayer Capsules.* Langmuir, **2000**, *16*, 1485.
3. Suttle, C.A., *Viruses in the sea.* Nature, **2005**, *437*, 356.
4. Lee, Y.J., et al., *Fabricating Genetically Engineered High-Power Lithium-Ion Batteries Using Multiple Virus Genes.* Science, **2009**, *324*, 1051.
5. Liepold, L., et al., *Viral capsids as MRI contrast agents.* Magn. Reson. Med., **2007**, *58*, 871.
6. Wu, W., et al., *Genome-Free Viral Capsids as Multivalent Carriers for Taxol Delivery.* Angew. Chem. Int. Edit., **2009**, *48*, 9493.
7. Douglas, T. and M. Young, *Host-guest encapsulation of materials by assembled virus protein cages.* Nature, **1998**, *393*, 152.
8. Minten, I.J., et al., *CCMV capsid formation induced by a functional negatively charged polymer.* Org. Biomol. Chem., **2009**, *7*, 4685.
9. Sikkema, F.D., et al., *Monodisperse polymer-virus hybrid nanoparticles.* Org. Biomol. Chem., **2007**, *5*, 54.
10. Comellas-Aragones, M., et al., *A virus-based single-enzyme nanoreactor.* Nat. Nanotechnol., **2007**, *2*, 635.
11. Minten, I.J., et al., *Controlled Encapsulation of Multiple Proteins in Virus Capsids.* J. Am. Chem. Soc., **2009**, *131*, 17771.
12. Fan, C.G., et al., *Short N-terminal sequences package proteins into bacterial microcompartments.* P. Natl. Acad. Sci. USA, **2010**, *107*, 7509.
13. Miletic, N., et al., *Immobilization of Candida antarctica lipase B on Polystyrene Nanoparticles.* Macromol. Rapid. Comm., **2009**, *31*, 71.
14. Zhang, N., et al., *Improving tolerance of Candida antarctica lipase B towards irreversible thermal inactivation through directed evolution.* Protein Eng., **2003**, *16*, 599.
15. Ying, Y. and L. Stefan, *Improved triglyceride transesterification by circular permuted Candida antarctica lipase B.* Biotech. Bioeng., **2009**, *105*, 44.
16. Chen, Q., H. Schonherr, and G.J. Vancso, *Block-Copolymer Vesicles as Nanoreactors for Enzymatic Reactions.* Small, **2009**, *5*, 1436.
17. Verduin, B.J.M., *Degradation of Cowpea Chlorotic Mottle Virus Ribonucleic-Acid Insitu.* J. Gen. Virol., **1978**, *39*, 131.
18. Verduin, B.J.M., *Preparation of CCMV-Protein in Connection with Its Association into a Spherical-Particle.* FEBS Lett., **1974**, *45*, 50.
19. Thesis of Claessen, V.I., Radboud University Nijmegen.

Summary

Viruses are increasingly utilized as nano building blocks. Their unparalleled ability to self-assemble into precisely defined nano-sized structures makes them uniquely suitable as nano-reactors or drug-delivery systems, and as novel materials. Nature has provided us with a multitude of virus types to choose from, each bearing their own distinct characteristics.

In our group, we are working with the cowpea chlorotic mottle virus (CCMV), an icosahedral RNA-containing plant virus with interesting assembly properties. It is possible to remove the RNA to obtain an empty virus shell, called capsid. The proteins forming the virus shell, i.e. capsid proteins, can be assembled and disassembled reversibly by adjusting the pH. Without the stabilizing interactions of the negatively charged RNA, the capsid disassembles at pH 7.5 into 90 capsid protein dimers. At pH 5.0, the capsid protein dimers re-assemble into the native 28 nm capsid.

Natural negatively charged polyelectrolytes like RNA are not the only compounds capable of templating the formation of capsids at high pH. It was previously shown by our group that polystyrene sulfonate (PSS), which is a negatively charged polymer, can also induce the formation of a capsid at pH 7.5. The obtained assembly, however, is smaller than that formed in the presence of RNA as a template at pH 5.0. In chapter 3, we describe studies aimed at obtaining more insight into the role of the polyelectrolyte in stabilizing the capsid, namely by the encapsulation of a negatively charged redox active polyelectrolyte, polyferrocenylsilane (PFS). This polyelectrolyte can be oxidized electrochemically, resulting in a polymer with a net neutral charge. Since the capsid is thought to be stabilized by the presence of negative charges, oxidation of the polyelectrolyte could remove these interactions, which might result in destabilization of the capsid. Although the results were inconclusive, it was shown that PFS can also induce the formation of a small 18 nm capsid.

In chapter 4, it is demonstrated that capsid-like structures not only assemble around linear polyelectrolytes at pH 7.5, but also around micelles formed by DNA amphiphiles. This provides a general method to load small molecules into the capsid either by hybridization to the DNA or by loading into the hydrophobic core

of the DNA particles. This could potentially have applications for the development of a drug delivery system.

To further explore and expand the properties of the CCMV capsid protein, it was heterologeously expressed in *E.coli* with the objective to perform genetic modifications. The production and purification of these genetically modified capsid proteins is discussed in chapter 5. They proved to be more resistant to degradation than the wild-type protein. The effect of the length of the N-terminus on the PSS-induced capsid formation is also described.

Previous studies conducted in our group by Dr. M. Comellas on the encapsulation of single enzymes inside the capsid had already provided new insights into the way enzymes function. The study of multiple enzymes in the confined space of the CCMV capsid was expected to provide greater understanding of the effect of confinement and local high concentration on enzyme activity. Confinement of enzymes is of interest since it mimics the natural cellular environment of enzymes, but the encapsulation of multiple enzymes is very difficult to achieve by statistical encapsulation. In chapter 6, the development of a new method, which increases the encapsulation efficiency and also allows control over the number of encapsulated enzyme is described s. To this end, an amino-acid motif called coiled-coil was attached to the capsid proteins, which could serve as a handle to attach proteins of interest. As a first test, enhanced green fluorescent protein (EGFP) was attached to the modified capsid proteins, and mixed with wild-type capsid protein prior to induction of capsid assembly. The number of encapsulated proteins was controlled by varying the ratio between the wild-type capsid protein and the capsid protein-EGFP complex. The assembly mechanism of this system appeared more complicated than expected and was analyzed in detail. These results are described in chapter 7.

Using the method described in chapter 6, multiple copies of a functional enzyme, Pseudozyma (formerly Candida) antarctica lipase B (PalB), were encapsulated inside the capsid. The effect of varying the amount of encapsulated PalB on the rate of substrate conversion was studied. The encapsulated PalB appeared to convert the substrate faster than the non-encapsulated PalB. Furthermore, the amount of encapsulated enzymes seemed to have little effect on the efficiency of the reaction. The reaction rates thus appear to be markedly influenced by the confinement and the resulting high local concentration of enzyme. These studies are described in chapter 9. The encapsulation of multiple different enzymes is important for the study of cascade reactions in the confined space of the

capsid. The method developed in chapter 6 also proved suitable to achieve co-encapsulation of EGFP and PalB, as reported in chapter 9.

The study of cascade reactions at pH 5.0 is challenging since most enzymes are not stable at low pH. This is a major drawback of the use of the CCMV capsid as a nanoreactor. Therefore, a new method was developed to stabilize the capsids at physiological pH. The N-terminus of the capsid protein was genetically modified with a histidine-tag, which is able to bind nickel ions. Capsids were stabilized at pH 7.5 by adding nickel ions to assembled capsids at pH 5.0, then increasing the pH to 7.5. By mixing EGFP-capsid protein complexes with the capsid proteins bearing a histidine tag, we successfully achieved encapsulation of EGFP at pH 7.5.

With the methods described in chapter 6, 8 and 9, it is now possible to encapsulate multiple different proteins and enzymes into the capsid in a controlled way, and to stabilize the capsid at physiological pH. It should now be possible to study cascade reactions inside the CCMV capsid. This might lead to exciting new insights and to the development of efficient nano-reactors.

In order to fully realize the potential of CCMV as nano-building block, it is essential to have a thorough understanding of the assembly properties of wild-type and modified capsid proteins, and to develop ways to manipulate and control their assembly. The studies described in chapters 3, 4, 5, 7 and 8 have contributed to this. Further development of the system proposed at the end of chapter 3 may elucidate the necessity of the negatively charged polymer for the stability of the capsid after assembly, by demolishing this polymer after it has induced the formation of the capsid.

The focus of this thesis is on the decoration of the capsid interior. This is only one of the three exploitable surfaces that CCMV has to offer. The exterior and interface of the capsid provide very different challenges and opportunities. The use of CCMV as a versatile nano-building block has just started. The present work provides several new methods that aid the utilization of CCMV as a biohybrid building block in nanotechnology.

Samenvatting

Virussen worden in toenemende mate gebruikt als bouwstenen voor de constructie van moleculaire systemen van nanometer afmeting. Ze zijn ongeëvenaard met betrekking tot het vermogen zichzelf spontaan te assembleren uit kleinere componenten, hetgeen ze erg geschikt maakt om gebruikt te worden in nano-reactoren, als medicijn-afgiftesystemen en als nieuwe materialen. De natuur heeft ons een grote diversiteit aan virussen geboden om mee te werken, elk met hun eigen karakteristieke eigenschappen.

In onze groep werken we met het cowpea chlorotic mottle virus (CCMV). Dit is een icosaedrisch plantenvirus, dat RNA bevat als genetisch materiaal, en interessant assembleergedrag vertoont. Het is mogelijk om het RNA te verwijderen, zodat een leeg omhulsel verkregen wordt, de eiwitmantel, die geassembleerd kan worden en weer uiteen kan vallen door de pH te veranderen. Zonder de stabiliserende interacties van het negatief geladen RNA, valt de mantel bij pH 7,5 uiteen in 180 manteleiwitten. Een verlaging van de pH naar 5,0 resulteert in het omgekeerde gedrag: 180 manteleiwitten vormen weer spontaan de eiwitmantel van 28 nm in doorsnede.

Negatief geladen polelektrolyten die voorkomen in de natuur, zoals RNA zijn niet de enige verbindingen die in staat zijn om als mal te dienen voor de formatie van de eiwitmantel bij hogere pH. In onze groep is eerder aangetoond dat een negatief geladen polymeer, polystyreensulfonaat (PSS), ook de vorming van de mantel kan induceren bij pH 7,5. De op deze manier verkregen structuur is echter kleiner dan wanneer de mantel zich vormt in het bijzijn van RNA als mal. In hoofdstuk 3 beschrijven we onderzoek dat gericht is op het verkrijgen van meer inzicht in de rol van het polymeer in de stabilisatie van de mantel, en wel door de insluiting van een negatief geladen redox-actief polyelektrolyt, polyferrocenylsilaan (PFS). Dit polyelektrolyt kan chemisch geoxideerd worden, hetgeen resulteert in een polymeer zonder lading. Omdat de mantel waarschijnlijk gestabiliseerd wordt door de aanwezigheid van negatieve ladingen, zou de interactie tussen het manteleiwit en het PFS dus kunnen verdwijnen door oxidatie van het polyelektrolyt, hetgeen mogelijk zou leiden tot destabilisatie van de mantel. Hoewel de resultaten

geen duidelijk uitsluitsel hierover gaven, werd wel aangetoond dat PFS ook de formatie van een klein 18 nm virusdeeltje kan induceren.

In hoofdstuk 4 wordt beschreven dat de manteleiwitten bij pH 7,5 niet alleen kunnen assembleren rondom lineaire polelektrolyten, maar ook rond micellen, die bestaan uit DNA-amphiphilen. Kleine moleculen kunnen aan de micellen bevestigd worden door middel van hybridisatie aan het DNA of door ze onder te brengen in de hydrofobe binnenkant van de micellen. Dus levert dit een algemene methode op om de binnenkant van de mantel te beladen met kleine moleculen. Dit zou mogelijk toepassingen kunnen hebben in de ontwikkeling van medicijn-afgiftesystemen.

Om de eigenschappen van de CCMV manteleiwitten verder te onderzoeken en uit te breiden, werden deze eiwitten tot expressie gebracht in E.coli-bacteriën om zo genetische modificaties aan te kunnen brengen. De productie en zuivering van deze genetisch gemodificeerde manteleiwitten wordt besproken in hoofdstuk 5. Ze bleken tevens beter bestand te zijn tegen degradatie dan het wild-type eiwit. Het effect van de lengte van de N-terminus op de door PSS geïnduceerde formatie van de mantel wordt ook in dit hoofdstuk beschreven.

Dr. M. Comellas heeft in onze groep onderzoek gedaan naar de insluiting van enkelvoudige enzymen in de virus mantel. Dit heeft nieuwe inzichten opgeleverd over de werking van enzymen. De verwachting was daarom dat de studie van meerdere enzymen in de beperkte ruimte van de CCMV mantel tot een beter begrip zou kunnen leiden van de effecten van een beperkte ruimte in combinatie met een hoge lokale enzym-concentratie op de activiteit van de enzymen. Insluiting van enzymen in kleine ruimtes is van belang omdat dit het gedrag in cellen nabootst, die de natuurlijke omgeving zijn van enzymen. De insluiting van meerdere enzymen in de CCMV mantel is echter erg moeilijk te realiseren met de statistische insluitingsmethode, die gebruikt is door Dr. M. Comellas. In hoofdstuk 6 wordt daarom de ontwikkeling van een nieuwe methode beschreven die de insluitingsefficiëntie verhoogt en die ook de mogelijkheid biedt om te regelen hoeveel enzymen er ingesloten worden. Om dit te bereiken werden de manteleiwitten genetisch gemodificeerd zodat ze een aminozuur-sequentie genaamd coiled-coil bevatten. Deze coiled-coil kon worden gebruikt als een soort klittenband om er andere eiwitten, uitgerust met de complementaire coiled-coil, aan te bevestigen. Als een eerste test werd een fluorescent eiwit, het "enhanced green fluorescent protein" (EGFP), bevestigd aan de gemodificeerde manteleiwitten, en vervolgens gemengd met wild-type manteleiwit, waarna de vorming van de mantel geïnduceerd werd door verlaging van de pH. Het aantal ingesloten eiwitten kon

gestuurd worden door de verhouding tussen het wild-type manteleiwit en het manteleiwit-EGFP complex te variëren. Het vormingsmechanisme van dit systeem bleek complexer te zijn dan gedacht en werd in detail geanalyseerd. De resultaten hiervan zijn beschreven in hoofdstuk 7.

Door de in hoofdstuk 6 ontwikkelde methode te gebruiken konden meerdere kopieën van een functioneel enzym, Pseudozyme (voorheen Candida) antarctica lipase B (PalB) ingesloten worden in de mantel van CCMV. Het effect van het variëren van het aantal ingesloten enzymen op de reactiesnelheid werd bestudeerd. Het bleek dat ingesloten PalB het substraat sneller omzette dan niet ingesloten PalB. Ook had de hoeveelheid ingesloten PalB enzymen weinig invloed op de efficiëntie van de reactie. De kleine ruimte en resulterende hoge concentratie van het enzym in deze ruimte bleek een grote invloed te hebben op de reactiesnelheid. Dit is beschreven in hoofdstuk 9. De insluiting van meerdere verschillende enzymen is belangrijk voor de studie van cascadereacties in de begrensde ruimte van de CCMV-mantel. De methode die beschreven staat in hoofdstuk 6 bleek geschikt te zijn om zowel PalB als EGFP aan de binnenkant van de eiwitmantel samen te brengen. Ook dit resultaat staat beschreven in hoofdstuk 9.

De studie van cascadereacties bij pH 5,0 is lastig omdat de meeste enzymen niet actief zijn bij lage pH. Dit is een groot nadeel van het gebruik van de CCMV eiwitmantel als nanoreactor. Een nieuwe methode werd daarom ontwikkeld om de mantel bij fysiologische pH te stabiliseren. De N-termini van de manteleiwitten werden door genetische modificatie voorzien van een histidine-tag, die in staat is om nikkelionen te binden. De mantel werd gestabiliseerd bij pH 7,5 door nikkelionen aan de mantel toe te voegen bij pH 5,0 en vervolgens de pH te verhogen naar 7,5. Door EGFP-manteleiwit-complexen te mengen met de manteleiwitten met een histidine-tag, kon EGFP ingesloten worden bij pH 7,5.

Met de methoden beschreven in hoofdstukken 6, 8 en 9 is het nu mogelijk om meerdere typen eiwitten en enzymen op een gecontroleerde manier in te sluiten, en om de mantel van het CCMV te stabiliseren bij fysiologische pH. Het zou daarom nu mogelijk moeten zijn om cascadereacties te bestuderen binnenin de CCMV-mantel. Dit zou kunnen leiden tot interessante nieuwe inzichten en de ontwikkeling van efficiëntere nanoreactoren.

Om het volledige potentieel van CCMV als nanobouwsteen te kunnen verwezenlijken is het essentieel om een grondig begrip te hebben van de assemblage eigenschappen van het wild-type en de gemodificeerde manteleiwitten, en om methoden te ontwikkelen om de assemblage-eigenschappen te kunnen sturen.

De studies beschreven in hoofdstuk 3, 4, 5, 7 en 8 hebben hieraan bijgedragen. Verdere ontwikkeling van een degradeerbaar polymeer template dat voorgesteld wordt aan het einde van hoofdstuk 3 zou licht kunnen werpen op de noodzaak van een negatief geladen polymeer voor de stabiliteit van de eiwitmantel nadat deze gevormd is. Dit zou bereikt kunnen worden door het polymeer te ontleden nadat het de vorming van de mantel geïnduceerd heeft.

Dit proefschrift heeft zich gericht op studies naar de modificatie van de binnenkant van de CCMV-eiwitmantel. Dit is echter maar één van de drie mogelijke oppervlaktes voor modificatie die CCMV te bieden heeft. De buitenkant en de tussenlaag van de eiwitmantel bieden weer hele andere uitdagingen en mogelijkheden. Het gebruik van CCMV als nanobouwsteen is pas net begonnen. In dit proefschrift werden verschillende nieuwe methoden beschreven die het toepassen van het CCMV als bouwsteen voor de constructie van biohybride systemen van nanometer-afmeting kunnen helpen.

List of abbreviations

1,2-PD	1,2-propanediol
AgNP	Silver nanoparticles
CalB	Candida antarctica lipase B
CCMV	Cowpea chlorotic mottle virus
cDNA	Copy DNA
CK	Capsid protein with K-coil
CMV	Cowpea mosaic virus
CP	Coat protein/capsid protein
C-terminus	Carboxyl-terminus of a protein or peptide
CV	Cyclic voltametry
DBCs	DNA block copolymers
DiFMU-octanoate	6,8-difluoro-4-methylumbelliferyl octanoate
DiI	1,1'-dioctadecyl- 3,3,3',3'-tetramethylindocarbocyanine perchlorate
DLS	Dynamic light scattering
DMSO	Dimethyl sulfoxide
DNA	Deoxyribonucleic acid
Dsp	Bacterial DNA binding proteins from starved cells
E1	Pyruvate decarboxylase component of PDH-complex
E2	Dihydrolipoyl acetyltransferase component of PDH-complex
E3	Dihydrolipoyl dehydrogenase component of PDH-complex
ECL	Enhanced chemiluminescence
E-coil	Negatively charged coiled-coil, amino acid sequence: $(EIAALEK_3)$
EDTA	Ethylenediaminetetraacetic acid
EDX	Energy dispersive X-ray
EGFP	Enhanced green fluorescent protein
E.coli	Escherichia coli
EM	Electron microscopy
Eut	Ethanolamine utilization
FCCS	Dual-color fluorescence cross-correlation spectroscopy
FPLC	Fast performance liquid chromatography
GE	EGFP with E-coil

GE-CK complex	Two GE proteins bound to one CK dimer
GFP	Green fluorescent protein
GST	Glutathion S-transferase
His-CP	Capsid protein with N-terminal His-tag
His-tag	Histidine-tag, an amino-acid sequence of multiple sequential histidines, in this thesis a sequence of 6 histidines was used
HRP	Horseradish peroxidase
HSFn	Horse spleen apoferritin
Hsp60	Heat shock protein 60
ICP-MS	Inductively coupled plasma-mass spectrometry
ICP-OES	Inductively coupled plasma-optical emission spectroscopy
IPTG	Isopropyl b-D-1-thiogalactopyranoside
K-coil	Positively charged coiled-coil, amino acid sequence: ($KIAALKE_3$)
K_d	Dissociation constant
LB	Lysogeny broth
MBP	Maltose binding protein
MCPs	Proteinaceous microcompartments
MPV	Murine polyoma virus
MVP	Major vault protein
MQ	Milli-Q water
MRI	Magnetic resonance imaging
MW	Molecular weight
MWCO	Molecular weight cut-off
NΔ34	The N in this notation indicates that the amino acids are missing from the N-terminus, Δ indicates missing amino acids, 34 states the number of missing amino acids
NHS	*N*-hydroxy succinimide
Ni-NTA	Nickel-nitrilotriacetic acid
NP	Nanoparticle
NTA	Nitrilotriacetic acid
N-terminus	Amino terminus of a protein or peptide
ODNs	Oligonucleotides
P11	DBCs containing PPO blocks with a molecular weight of 6800 g/mol, attached to DNA of11 nucleotides in length
P22	DBCs containing PPO blocks with a molecular weight of 6800 g/mol, attached to DNA of 22 nucleotides in length
PalB	Pseudozyma (formerly Candida) antarctica lipase B
PCB-1	Paramecium bursaria chlorella virus type I
PCR	Polymerase chain reaction

PDH	Pyruvate dehydrogenase
Pdu	Propanediol utilization
PduP	CoA-dependent propionaldehyde dehydrogenase
PEG	Polyethylene glycol
PEI	Poly(ethylene)imine
PfFt	Pyrococcus furiosus ferritin
PFS	Polyferrocenyl silane
PPO	Polypropylene oxide
PSBD	Peripheral subunit-binding domain of the E2 component of the PDH-complex
PSS	Polystyrene sulfonate
PVX	Potato virus X
RNA	Ribonucleic acid
ROX	6-carboxylic-X-rhodamine
RT	Room temperature
RuBisCO	Ribulose bis-phosphate carboxylase
SDS-PAGE	Sodium dodecyl sulfate-poly acrylamide gel electroforese
SFM	Scanning force microscopy
sHsp	Small heat shock protein
STIV	Sulfolobus turreted icosahedral virus
STMV	Satellite tobacco mosaic virus
SV40	Simian virus 40
SWNT	Single-walled carbon nanotubes
tcHis-CP	Thrombin cleaved His-CP
TEM	Transmission electron microscopy
TEP1	Telomerase-associated protein 1
TMV	Tobacco mosaic virus
Tris	Tris(hydroxymethyl)aminomethane
UC	Unknown complex
UU11	Lipid-DNA 11-mer containing two 5-dodec-1-ynyluracil nucleobases at the 5'- end followed by 9 naturally occuring nucleobases
UV-Vis	Ultraviolet-visual
VLP	Virus like particle
VP1	Virus protein 1
VPARP	Vault poly(ADP ribose polymerase)
wt CP	Wild-type capsid protein
wt CP NΔ34	Truncated capsid protein lacking the first 34 amino acids of the N-terminus
yCD	Yeast cytosine deaminase

Dankwoord

Dit dankwoord geeft me de kans om iedereen die op de een of andere manier aan dit proefschrift bijgedragen heeft te bedanken, en daar maak ik natuurlijk graag gebruik van. Waar we normaal gesproken in artikelen en presentaties alleen de namen noemen van degenen die een tastbare wetenschappelijke bijdrage hebben geleverd, krijg ik hier de gelegenheid om iedereen wat uitgebreider te bedanken, inclusief al degenen die op andere, maar minstens zo belangrijke manieren, hebben bijgedragen aan de totstandkoming van dit proefschrift.

Roeland en Jeroen, bedankt voor de grote vrijheid die ik gekregen heb bij het verkennen van het geweldige virus onderzoeksonderwerp. Jullie vertrouwen en vermogen om overal een positieve draai aan te geven, maakten dat ik na een gesprek met jullie altijd weer vol frisse moed buiten stond.

I would like express my gratitude to the manuscript committee, Jan van Hest, David Reinhoudt and Trevor Douglas, for reading and correcting my thesis.

Yujie, it was a great pleasure to work (and chat) with you. The research described in chapter 3 would not have been possible without you. It is funny how many times we kept bumping into each other after that.

Minseok, you were the driving force behind the research in chapter 4, you always wanted to go faster in your enthusiasm. The funny micelle puppets you designed were the direct inspiration for the carpenter on my cover.

Alan and Kerstin, without your enthusiasm and feisty discussions about the research in chapter 9, we might still not know how to interpret the results. Thank you for correcting the article so quickly and thoroughly.

Ik heb de afgelopen jaren meerdere malen gezien dat de confocale laser geen apparaat is waar je 'even' wat metingen op doet. Jammer genoeg strooide de vulkaan in IJsland ook nog eens roet in het eten. Arend, bedankt dat je geprobeerd hebt de metingen toch door te laten gaan. Victor, dank je voor de inzet om direct na terugkomst uit Amerika achter de confocal te duiken. Het heeft de feitelijke onderbouwing geleverd enkele belangrijke aannames in hoofdstuk 9.

Naar huidige maatstaven heb ik geloof ik niet veel studenten gehad, maar liever kwaliteit dan kwantiteit, en dus heb ik absoluut niets te klagen. Koos, je hield van discussiëren, en we hebben dan ook vele discussies gevoerd over jouw (en

mijn) onderzoek, wat zich daar gelukkig ook uitstekend voor leende. Het heeft geresulteerd in een fantastisch stukje onderzoek, waar hoofdstuk 8 volledig aan gewijd is. Saskia, nadat Koos weg was, was jij er gelukkig nog wel, zodat ik ook tijdens het laatste halve jaar van mijn promotie iemand had om gezellig mee te kletsen. Je hebt veel synthese gedaan in je project. De literatuur op dit gebied bleek helaas niet helemaal betrouwbaar, maar dankzij jou doorzettingsvermogen en de input van Rene Aben en Hans Scheeren is het toch nog gelukt om het monomeer te maken. De scriptie die je geschreven hebt, heeft als basis gediend voor mijn tweede hoofdstuk en een review artikel, en omdat je dat zo'n geweldig goed werk af hebt geleverd, heb ik daar nauwelijks iets aan hoeven te veranderen.

When I started my research in the 'virusgroup' Marta, Friso and Linda had already laid de groundworks for my research. Marta and Friso, not only was your work the inspiration for a large part of this thesis, you also taught me a lot about the virus in general and how to extract it from the plants. Linda, jij hebt me niet alleen van het andere belangrijke basismateriaal voorzien; het plasmide wat codeert voor het bacteriële capsid eiwit, maar jij bent ook degene geweest die me in het begin wegwijs heeft gemaakt op het lab, en me zoveel verschillende technieken heeft geleerd. Ontzettend bedankt!

Guillaume, Andrès, Mauri and Christine, it was really nice having you here, growing plants in increasingly obscure corners of the university wouldn't have been half as much fun without all of you.

All the other people of the Nolte group, thanks for the great time in and around the lab.

By the time this thesis is printed, the virus group in Nijmegen will have almost ceased to exist, but luckily a new and enthusiastic group of scientists are taking over the virus research in Jeroens new group in Twente. Rik, Melanie, Anne, Martijn, Oya and Yujie (again), I have enjoyed getting to know you and being able to help a little bit. It is nice to know that some of my projects are continued, and I'm sure with you the research is in very capable hands.

Contrary to popular believe, I was never officially a part of the van Hest group (at least not during my PhD), but as the Nolte group was declining in size, they were kind enough to more or less adopt us, and I want to thank the entire van Hest group for the 'gezelligheid' during lunchtime, drinks, and other occasions. Hans, bedankt voor de hulp met peptide synthese en het delen van een schijnbaar onuitputtelijke hoeveelheid verhalen. Joris, het is moeilijk om jou ergens specifiek voor te bedanken, want je hielp, gevraagd of ongevraagd, altijd met alles.

Linda, Sanne, Rosalie, Mark, Mark en Mark (tja, ik kan er ook niets aan doen, het is blijkbaar een populaire naam), het was fijn om een paar collega's te hebben bij wie ik te rade kon gaan met moleculair biologische problemen. Bedankt ook voor alle gezelligheid aan de soms wat stille 'overkant'.

Behalve op het NCMLS heb ik ook veel tijd besteed in het algemeen instrumentarium, waar Liesbeth, Jelle, Geert-Jan en Rien altijd klaar stonden om te helpen. Iedere keer dat de FPLC weer iets vreemds deed dacht ik weer "de volgende keer heb ik Rien niet nodig, want nu heb ik toch wel alles gezien wat er mis kan zijn met dit apparaat", en elke keer had ik het weer mis. Rien ontzettend bedankt voor al je hulp!

Peter, ondanks de bestelsystemen heb ik dankzij jou de chemicallieen vaak toch nog binnen een redelijke tijd gekregen, dank je voor alle extra moeite. Desiree, zonder jou zouden vele dingen minder soepel zijn verlopen. Dank je voor al je hulp.

De afgelopen jaren ben ik twee keer naar het buitenland geweest voor een conferentie. Beide keren een fantastische ervaring, niet in het minst dankzij Sanne, Morten, Maaike, Stijn en Dennis, met wie ik Oxford verkend heb, en TuHa, Victor, Luiz en Loes, waarmee ik door de steile straten van San Francisco heb gewandeld.

Stijn en Loes, fijn dat twee vrienden die ik al ken vanaf de allereerste dag hier in Nijmegen mijn paranimfen willen zijn. Samen met Willem, Nearchos, Arjan, Vivike en Otmar, hebben we heel wat gezellige dagen en avonden doorgebracht, de afgelopen jaren zouden lang niet zo leuk geweest zijn zonder jullie.

En dan is er natuurlijk nog het thuisfront. Jan en Myriam, jullie hebben me van kinds af aan enthousiast gesteund in alles wat ik deed, en daar zijn jullie nooit mee opgehouden, dankzij jullie ben ik gezegend met een extra paar 'grootouders'.

Peter, mijn 'kleine' broertje, je hebt de afgelopen jaren enorme sprongen voorwaarts gemaakt, en me keer op keer versteld doen staan. Ik ben enorm trots op je.

Pap en mam, jullie hebben altijd iedere kans, die jullie zagen om me te helpen, aangegrepen, bedankt voor de onvoorwaardelijke steun tijdens mijn hele leven.

List of publications

Metal ion induced formation and stabilization of protein cages based on the cowpea chlorotic mottle virus.
Inge J. Minten, Koos D.M. Wilke, Linda J.A. Hendriks, Jan C.M. van Hest, Roeland J.M. Nolte, Jeroen J.L.M. Cornelissen
Submitted for publication

Catalytic capsids: the art of confinement.
Inge J. Minten, Victor I. Claessen, Kerstin Blank, Alan E. Rowan, Roeland J.M. Nolte, Jeroen J.L.M. Cornelissen
Chem. Sci., DOI: 10.1039/c0sc00407c.

Virus-like particles templated by DNA micelles: a general method for loading virus nanocarriers.
Minseok Kwak, Inge J. Minten, Diana-Milena Anaya, Andrew J Musser, Melanie Brasch, Roeland J.M. Nolte, Klaus Müllen, Jeroen J.L.M Cornelissen, Andreas Herrmann
J. Am. Chem. Soc., **2010,** 132, 7834.

Complex assembly behavior during the encapsulation of green fluorescent protein analogs in virus derived protein capsules.
Inge J. Minten, Roeland J.M. Nolte, Jeroen J.L.M. Cornelissen
Macromol. Biosci., **2010,** 10, 539.

Controlled encapsulation of multiple proteins in virus capsids.
Inge J. Minten, Linda J.A. Hendriks, Roeland J.M. Nolte, Jeroen J.L.M. Cornelissen
J. Am. Chem. Soc., **2009,** 131, 17771.

CCMV capsid formation induced by a functional negatively charged polymer.
Inge J. Minten, Yujie Ma, Mark A. Hempenius, G. Julius Vancso, Roeland J.M. Nolte, Jeroen J.L.M. Cornelissen
Org. Biomol. Chem., **2009,** 7, 4685.

Controlled disassembly of peptide amphiphile fibres.
Dennis W.P.M. Löwik, Joris T. Meijer, Inge J. Minten, Henri Van Kalkeren, Lisa Heckenmuller, Ines Schulten, Kwinten Sliepen, Peter Smittenaar, Jan C.M. van Hest
J. Pept. Sci., **2008,** 14, 127.

Disassembling peptide-based fibres by switching the hydrophobic-hydrophilic balance.
Joris T. Meijer, Marloes J. A. G. Henckens, Inge J. Minten, Dennis W.P.M. Löwik, Jan C.M. van Hest
Soft Matter, **2007,** 3, 1135.

Curriculum vitae

Inge Minten werd op 11 mei 1983 geboren in Sterksel. Aan het Were Di college in Valkenswaard haalde zij in 2001 haar VWO-diploma. In hetzelfde jaar begon ze met de studie moleculaire levenswetenschappen in Nijmegen. Deze studie behelsde twee stages. De eerste werd uitgevoerd in de groep van Professor Jan van Hest met als onderwerp de synthese van een photocleavable linker gekoppeld aan een cell penetrating peptide voor de constructie van een getarget drug-delivery systeem. De tweede stage werd uitgevoerd in de Molecular Design and Informatics groep van Professor Jacob de Vlieg bij Organon in Oss. Na het afsluiten van haar studie in 2006 is Inge begonnen met een promotieonderzoek in de groep van Prof. Roeland Nolte en Prof. Jeroen Cornelissen. De resultaten van dit onderzoek staan beschreven in dit proefschrift.